

Über die mathematische Wahrscheinlichkeit

Autor(en): **Finsler, P.**

Objektyp: **Article**

Zeitschrift: **Elemente der Mathematik**

Band (Jahr): **2 (1947)**

Heft 6

PDF erstellt am: **30.06.2024**

Persistenter Link: <https://doi.org/10.5169/seals-12828>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Über die mathematische Wahrscheinlichkeit¹⁾

Man hat sich lange Zeit mit einer sehr einfachen und anschaulichen Definition der mathematischen Wahrscheinlichkeit begnügt, mit der *klassischen Definition*, nach welcher die Wahrscheinlichkeit w für ein Ereignis gleich dem Quotienten der Anzahl g der für das Ereignis günstigen Fälle dividiert durch die Anzahl m aller möglichen Fälle ist, vorausgesetzt, daß diese Fälle sich ausschließen und gleichmöglich sind: $w = \frac{g}{m}$. Gegen diese Definition sind dann aber Einwände erhoben worden, die ich noch erwähnen werde. Ob oder wie weit diese Einwände berechtigt sind, wird auch noch zu untersuchen sein.

Man hat dann nach einer neuen Definition der Wahrscheinlichkeit gesucht, und viele glaubten, daß in einer Häufigkeits- oder, besser gesagt, *Limesdefinition* eine exakte Begründung der Wahrscheinlichkeitsrechnung gefunden worden sei. Man setzte die Wahrscheinlichkeit gleich dem Grenzwert der relativen Häufigkeit, wenn der betreffende Versuch unendlich oft wiederholt wird, also $w = \lim_{n \rightarrow \infty} \frac{a_n}{n}$, wenn bei den n ersten Versuchen das erwartete Ereignis a_n -mal eintritt und n gegen unendlich geht. Aber auch gegen diese Definition sind triftige Einwände möglich, insbesondere was die Anwendbarkeit in der Praxis betrifft, da man eben einen Versuch praktisch nicht unendlich oft ausführen kann.

Schließlich ist man dann wieder auf eine Methode zurückgekommen, die schon vorher versucht worden war, nämlich auf die *axiomatische Methode*, die sich ja z. B. in der Geometrie so gut bewährt hat. Die Schwierigkeit liegt hier darin, die richtigen Axiome zu finden, also einfache und plausible Annahmen für die Wahrscheinlichkeit, aus denen sich dann alles übrige ergeben soll.

Wenn man nun die Axiomensysteme, wie sie in neuerer Zeit aufgestellt und verwendet wurden, näher untersucht, so findet man, daß sie eigentlich von der klassischen Definition gar nicht so weit entfernt sind, daß sie im wesentlichen nur in etwas anderer Form dasselbe besagen. Damit sind sie aber auch ähnlichen Einwänden ausgesetzt, und sie bedürfen in der Tat noch einer Ergänzung, um als ausreichend betrachtet zu werden.

Ich glaube, daß für eine Einführung in die Wahrscheinlichkeitsrechnung und deshalb auch für den Schulunterricht die klassische Methode immer noch den besten Weg darstellt und daß sie nur noch etwas ergänzt werden muß, um auch mathematisch einwandfrei zu sein.

Ehe ich diese Dinge näher ausführe, muß ich zuerst noch die Aufgabe und den Gegenstand der Wahrscheinlichkeitsrechnung genauer festlegen.

Man hat es in der Wahrscheinlichkeitsrechnung mit *Versuchen* zu tun, z. B. dem Werfen eines Würfels, dem Ziehen aus einer Urne, der Ausführung einer Beobachtung oder Messung usw. Das Ergebnis eines solchen Versuchs nennen wir ein *Ereignis*; so

¹⁾ Vortrag an der 51. Jahresversammlung des Vereins Schweizerischer Mathematiklehrer in Baden, Oktober 1947.

z. B. das Erscheinen einer 6 beim Würfeln. Die Versuchsbedingungen oder die *Voraussetzungen* des Versuchs sollen gegeben sein, aber nicht so vollständig, daß dadurch das Ergebnis eindeutig bestimmt wäre. Es wird z. B. gesagt, daß mit einem bestimmten Würfel geworfen werden soll; es wird aber nichts gesagt über die Ausgangslage und Ausgangsgeschwindigkeit, und dies hat zur Folge, daß unter denselben Voraussetzungen verschiedene Ergebnisse möglich sind.

Es wird nun weiter angenommen, daß bei gegebenen Versuchsbedingungen jedem von diesen möglichen Ereignissen eine bestimmte *Wahrscheinlichkeit* zukommt, so daß man also sagen kann, aus den Voraussetzungen V folgt das Ereignis E mit der Wahrscheinlichkeit $w_V(E)$:

$$V \rightarrow E \quad \text{mit} \quad w = w_V(E).$$

Die Voraussetzungen sollen also, soweit nötig, als Index angeschrieben werden.

Die Aufgabe der Wahrscheinlichkeitsrechnung besteht nun darin, diese Wahrscheinlichkeit zu definieren und dann im Einzelfall zu berechnen.

Bei der klassischen Definition wird die Annahme gemacht, daß die bei gegebenen Voraussetzungen möglichen Ereignisse entweder gleichmöglich sind oder sich wenigstens in gleichmögliche Ereignisse zerlegen lassen. Mit diesem Zusatz wird schon einem Einwand begegnet, der gelegentlich erhoben wurde. Es hieß, daß man mit der klassischen Definition wohl den symmetrischen, idealen Würfel behandeln könne, bei dem alle Seiten gleichmöglich sind, aber nicht einen gewöhnlichen, realen Würfel, der gewisse Unsymmetrien aufweist. Aber man braucht ja nicht nur auf die Endlage des Würfels zu achten, wenn er zur Ruhe gekommen ist, sondern kann auch die verschiedenen Möglichkeiten ins Auge fassen, wie er auf dem Tisch auftritt, in welcher Lage und mit welcher Geschwindigkeit, und kann sich diese Fälle in gleichmögliche aufgeteilt denken, von denen dann ein Teil für eine bestimmte Endlage, z. B. für eine 6, günstig ist. Bei dieser Auffassung müssen dann aber die Wahrscheinlichkeiten für die verschiedenen Seiten nicht mehr einander gleich sein. Durch einen Grenzübergang kann man dann auch unendlich viele Fälle und auch irrationale Wahrscheinlichkeiten berücksichtigen; das macht keine Schwierigkeit.

Der Haupteinwand, der gegen die klassische Definition erhoben wurde, ist aber ein anderer. Es wird behauptet, die Definition enthalte einen fehlerhaften Zirkel, denn sie stütze sich auf die Voraussetzung von gleichmöglichen, das heißt aber von gleichwahrscheinlichen Fällen; die Wahrscheinlichkeit sei also erst definiert, wenn man schon wisse, was gleichwahrscheinlich bedeutet. Dieser Einwand ist nicht berechtigt. Die gleichmöglichen Fälle ergeben sich nach der Definition auch als gleichwahrscheinlich, das ist ganz in Ordnung. Im übrigen aber weist der Einwand nur auf eine wichtige Eigenschaft der Wahrscheinlichkeit hin, die beachtet werden muß: Abgesehen von den Grenzfällen der Sicherheit und der Unmöglichkeit kann nämlich die Wahrscheinlichkeit nie, auch nicht durch beliebig viele Beobachtungen, in absolutem Sinne festgelegt werden, ohne daß *a priori* eine *Annahme* gemacht wird über andere Wahrscheinlichkeiten oder über die Gleichheit von solchen. Das ist aber kein Fehler, so wenig wie es in der Geometrie ein Fehler ist, daß die Länge einer Strecke nicht in absolutem Sinn festgelegt wird, sondern erst, wenn wenigstens eine Länge gegeben ist und auch über die Gleichheit von Längen Festsetzungen getroffen werden. In der Geometrie benutzt man den Zirkel, um theoretisch oder praktisch gleiche Längen

zu konstruieren; in der Wahrscheinlichkeitsrechnung hat man den Würfel oder die Urne, um theoretisch oder praktisch gleiche Wahrscheinlichkeiten zu konstruieren. In der Wahrscheinlichkeitsrechnung kann schon die Gleichheit von Wahrscheinlichkeiten genügen, um diese vollständig festzulegen: Wenn die 6 Seiten eines Würfels gleichwahrscheinlich sind, dann hat jede die Wahrscheinlichkeit $1:6$. Mindestens eine solche Annahme über Gleichwahrscheinlichkeit ist aber notwendig, wenn man für die Wahrscheinlichkeiten bestimmte Zahlen erhalten will. Diese Annahmen sind in die Voraussetzungen V aufzunehmen.

Ein anderer Einwand gegen die klassische Definition ist zum Teil berechtigt, nämlich der, daß die Definition keine Auskunft darüber gibt, welche Fälle als gleichmöglich zu betrachten sind. Ich sage «zum Teil», denn zunächst ist es ja, wie eben bemerkt, eine reine Festsetzung, welche Fälle als gleichmöglich gelten sollen. Man kann z. B. fordern, daß die Seiten eines gedachten Würfels gleichmöglich sein sollen, man kann aber auch einen nichtidealen Würfel annehmen, dessen Seiten verschiedene Wahrscheinlichkeiten haben, genau so, wie man in der Geometrie ein Dreieck als gleichseitig oder auch als ungleichseitig voraussetzen kann. Auch wenn man ein *Symmetrieprinzip* verwendet, nach dem solche Fälle, die bis auf die Bezeichnung in gleicher Weise in die Voraussetzungen eingehen, auch gleiche Wahrscheinlichkeiten erhalten sollen, so ist auch das eine willkürliche, aber unter Umständen zweckmäßige Festsetzung

Aber in einem Falle sollte diese Festsetzung nicht mehr willkürlich sein, nämlich dann, wenn es sich um die *Kombination* von Ereignissen handelt. Wenn ich z. B. mit einem idealen Würfel zweimal werfe, so sollte ich schließen können, daß alle 36 Kombinationen $1, 1$ bis $6, 6$ gleichmöglich sind und also jede die Wahrscheinlichkeit $1:36$ erhält. Dies folgt aber nicht aus der Definition und auch nicht aus dem Symmetrieprinzip, denn es ist etwas anderes, ob eine Seite sich wiederholt oder ob sie wechselt. Mit andern Worten, der *Multiplikationssatz* der Wahrscheinlichkeitsrechnung läßt sich aus der gegebenen Definition allein nicht herleiten, und das ist eine Lücke, die ausgefüllt werden muß.

In der Limesdefinition wird die Wahrscheinlichkeit als Grenzwert der relativen Häufigkeit eingeführt, wenn die Anzahl der Versuche ins Unendliche geht. Da man aber die Versuche doch nicht unendlich oft ausführen kann, so setzt man als Näherungswert für die Wahrscheinlichkeit die relative Häufigkeit bei einer hinreichend langen Versuchsreihe. Ist dies berechtigt? Und was heißt hier «hinreichend lang»? Eine unendliche Folge durch eine endliche zu ersetzen ist jedenfalls dann, aber auch nur dann berechtigt, wenn man den Fehler abschätzen kann und sieht, daß er nicht zu groß ist. Wenn man sich auf den Standpunkt der klassischen Theorie stellt, so kann man das Bernoullische Theorem anwenden und damit eine solche Abschätzung vornehmen. Mit der Limesdefinition geht dies aber nicht, denn hier hat man gar keinen Zusammenhang zwischen der endlichen Folge und der unendlichen. Ich will dies an einem Beispiel zeigen:

Es sei die Aufgabe gestellt, die Wahrscheinlichkeiten für die Seiten eines bestimmten realen Würfels durch Versuche zu ermitteln. Wie oft muß man werfen? Bei 100 oder 1000 Würfeln wird man noch ziemlich grobe Werte erhalten, bei 100000 Würfeln jedenfalls schon so gute, daß die Unsymmetrie des Würfels erkennbar wird. Darf man nun hier abbrechen und diese Werte als Näherungen für die Grenzwerte betrach-

ten? Um dies zu prüfen, müßte man die Versuche noch weiter fortsetzen. Man könnte sich etwa eine Maschine denken, welche dies ausführt und vielleicht millionen- und millardenmal würfelt und die Ergebnisse notiert. Was wird dabei herauskommen? Nun, der zu untersuchende Würfel wird sich bei dieser Inanspruchnahme so sehr abnützen, daß die schließlichen Grenzwerte, wenn man überhaupt von solchen reden kann, ganz anders ausfallen werden als die ersten Näherungswerte. Tatsächlich hat eben der Grenzwert bei einer unendlichen Folge mit der Häufigkeit bei einer bestimmten endlichen Folge gar nichts zu tun, wenn nicht noch besondere Bedingungen hinzukommen, die hier aber fehlen. Dies ist ein Einwand praktischer Natur; die Limestheorie läßt sich in der Praxis nicht anwenden.

Ich möchte noch einen Einwand theoretischer Natur erwähnen. Bei einem idealen Würfel ist es nach Voraussetzung möglich, daß er beim ersten Wurf gerade eine 6 zeigt. Auch beim zweiten Wurf ist dies möglich, ebenso beim dritten usw. Es ist also auch möglich, daß bei jedem Wurf immer eine 6 erscheint. Dies liefert aber den Grenzwert 1 und nicht $1:6$ und muß deshalb in der Limestheorie ausgeschlossen werden. Es wird also ein möglicher Fall als unmöglich erklärt; das ist ein Widerspruch, das sollte nicht sein.

Schließlich ist noch folgendes zu bemerken: Soweit man das Ergebnis eines Versuchs oder einer Versuchsreihe schon kennt, hat man Sicherheit und braucht keine Wahrscheinlichkeit. Wenn man aber das Ergebnis nicht kennt, dann kennt man auch nicht die relative Häufigkeit und noch weniger ihren Grenzwert, dann kennt man also nach dieser Definition auch keine Wahrscheinlichkeit.

Ich komme nun zur axiomatischen Grundlegung der Wahrscheinlichkeitsrechnung, wie sie etwa bei KOLMOGOROFF¹⁾ und neuerdings bei CRAMÉR²⁾ dargestellt ist. Es werden hier Axiome oder Postulate aufgestellt, denen die Wahrscheinlichkeit genügen soll.

Als erstes wird die *Existenz* der Wahrscheinlichkeit postuliert, und zwar soll bei gegebenen Versuchsbedingungen in einem bestimmten Umfang jeder *Menge* von möglichen Ereignissen eine Wahrscheinlichkeit zugeordnet sein, nämlich die Wahrscheinlichkeit dafür, daß das eintreffende Ereignis gerade dieser Menge angehört. Es soll also eine Mengenfunktion $w(\Sigma)$ existieren, die auch als $w(E \in \Sigma)$ geschrieben werden kann. Beim Würfeln kann Σ z. B. die Menge der Zahlen 2, 4 und 6 bedeuten; $w(\Sigma)$ ist dann die Wahrscheinlichkeit dafür, daß 2 oder 4 oder 6, das heißt, daß eine gerade Zahl erscheint.

Das zweite Axiom gibt die *Normierung* dieser Funktion: Die Wahrscheinlichkeit $w(\Sigma)$ soll eine *nichtnegative, additive* Mengenfunktion sein mit $w(\Omega) = 1$, wenn Ω die Menge aller möglichen Ereignisse bedeutet. Es wird also verlangt:

1. $w(\Sigma) \geq 0$,
2. $w(\Sigma_1 + \Sigma_2) = w(\Sigma_1) + w(\Sigma_2)$ für $\Sigma_1 \Sigma_2 = 0$,
3. $w(\Omega) = 1$.

$\Sigma_1 \Sigma_2 = 0$ bedeutet, daß die beiden Mengen Σ_1 und Σ_2 sich ausschließen, das heißt kein Ereignis gemeinsam haben; 2. ist das bekannte Additionstheorem für sich ausschließende Ereignisse; 3. bedeutet, daß einem sicheren Ereignis die Wahrscheinlichkeit 1 zukommen soll.

¹⁾ A. KOLMOGOROFF, Grundbegriffe der Wahrscheinlichkeitsrechnung. Berlin 1933.

²⁾ H. CRAMÉR, Mathematical Methods of Statistics. Princeton 1946.

Schließlich wird in einem dritten Axiom noch verlangt, daß die *Kombination* von Versuchen wieder einen Versuch ergibt, so daß dann nach dem ersten Axiom wieder eine zugehörige Wahrscheinlichkeit existiert.

Man kann sich nun das zweite Axiom noch näher veranschaulichen: Man denke sich etwa die möglichen Ereignisse als Punkte auf einer Achse aufgetragen; sie können diskret oder kontinuierlich angeordnet sein. Die Wahrscheinlichkeit w werde als *Masse* gedeutet, welche auf diese Ereignisse verteilt wird; sie ist nichtnegativ, und die Gesamtmasse soll gleich 1 sein. Man sieht leicht, daß dann auch das Additionstheorem erfüllt ist.

Wenn man nun weiter annimmt, daß diese Masse atomistisch aufgebaut, also aus lauter gleichschweren Atomen zusammengesetzt ist, und wenn man dann diese Atome als Repräsentanten gleichmöglicher Fälle betrachtet, so kommt man zu der klassischen Definition zurück: Die Wahrscheinlichkeit ist das Verhältnis der günstigen Fälle zu allen gleichmöglichen.

Wie nun aber diese Masse im einzelnen verteilt werden soll, darüber gibt das Axiomensystem keine Auskunft, und damit hat man auch denselben Einwand wie bei der klassischen Definition. Zunächst kann zwar die Verteilung für die Ausgangsereignisse willkürlich festgelegt werden, aber für die Kombination von unabhängigen Ereignissen sollte sie sich dann berechnen lassen. Das dritte Axiom liefert mit dem ersten zusammen nur die Existenz einer zugehörigen Wahrscheinlichkeit, aber nicht ihre Größe. Dieser Punkt muß also noch näher untersucht werden.

Zwei Ereignisse A und B heißen voneinander *unabhängig*, wenn die Wahrscheinlichkeit für das eine nicht vom Eintreffen oder Nichteintreffen des andern abhängt. Insbesondere ist B von A unabhängig, wenn $w_A(B) = w_{\bar{A}}(B) = w(B)$ ist, wobei \bar{A} das Nichteintreffen von A bedeuten soll. Für solche unabhängigen Ereignisse soll nun das Multiplikationstheorem gelten:

$$w(A, B) = w(A) w(B).$$

Wie schon bemerkt, folgt dies nicht aus den bisherigen Axiomen.

In der erwähnten Literatur wird nun einfach definiert: Die Ereignisse A und B heißen voneinander unabhängig, wenn für sie das Multiplikationstheorem erfüllt ist. Aber die Unabhängigkeit hat zunächst eine andere Bedeutung, und man kann nicht einen fehlenden Beweis durch eine Definition ersetzen. Man könnte sich so helfen, daß man den Satz als Axiom formuliert, dann braucht man ihn nicht zu beweisen. Es wäre dies aber ein recht kompliziertes Axiom, das nicht unmittelbar einleuchtet. Als Axiome sollen doch nur einfache und anschauliche Beziehungen genommen werden, aus denen sich die komplizierteren Sätze herleiten lassen. Gibt es eine einfache Aussage, aus der dieser Satz folgt? Das ist in der Tat der Fall, wenn man sich an die Grundauffassung der klassischen Theorie hält, das heißt an die Annahme, daß sich die möglichen Ereignisse in gleichmögliche zerlegen lassen. Es genügt dann, das folgende Axiom anzunehmen:

Gleichmögliche Ereignisse bleiben gleichmöglich, wenn eine für sie sicher erfüllte Bedingung vorausgesetzt wird.

Ich will das noch erläutern. Die Wahrscheinlichkeit eines Ereignisses wird sich im allgemeinen ändern, wenn eine neue Bedingung vorausgesetzt wird, auch wenn diese Bedingung für das Ereignis sicher erfüllt ist. Wenn man z. B. beim Würfeln nur

die geradzahigen Ergebnisse beachtet und die andern wegläßt, so hat man die Bedingung hinzugefügt, daß das Resultat eine gerade Zahl sein soll. Die Wahrscheinlichkeit für die Zahl 6 ist dann nicht mehr 1:6, sondern 1:3, denn man hat einen günstigen Fall unter den drei gleichmöglichen Fällen 2, 4 und 6. Warum sind diese Fälle gleichmöglich? Nur deshalb, weil sie ursprünglich gleichmöglich sein sollten und die hinzugefügte Bedingung, eine gerade Zahl zu sein, für sie sicher erfüllt ist; hier wird also das Axiom schon verwendet.

Allgemein läßt sich dieses Axiom der *Aussonderung* so formulieren:

$$w(A, B) = w(A, C) \rightarrow w_A(B) = w_A(C),$$

denn für die Ereignisse (A, B) bzw. (A, C) ist die Bedingung A sicher erfüllt und $w_A(B)$ bedeutet dasselbe wie $w_A(A, B)$. Es besagt: Wenn man diejenigen Ereignisse aussondert, für welche A sicher erfüllt ist, so soll unter diesen zwar nicht die Wahrscheinlichkeit, wohl aber die Gleichwahrscheinlichkeit erhalten bleiben.

Das Axiom findet sich im wesentlichen bei MARKOFF¹⁾, aber in einer Formulierung, die zu Fehlern Anlaß geben kann. Es ist auch nur dann ausreichend, wenn man noch das Existenzaxiom zur Verfügung hat.

Nunmehr läßt sich der Multiplikationssatz beweisen, und zwar in der allgemeinen Form, in der die Ereignisse nicht unabhängig sein müssen:

Sind A und B zwei beliebige Ereignisse, so ist es möglich, daß sie beide eintreffen, oder nur eines, oder keines. Da diesen Fällen nach dem ersten und dritten Axiom bestimmte Wahrscheinlichkeiten zukommen, kann man annehmen, daß sie sich in gleichmögliche Fälle zerlegen lassen. α davon seien günstig für (A, B) , β für (A, \bar{B}) , γ für (\bar{A}, B) und δ für (\bar{A}, \bar{B}) . $\alpha + \beta + \gamma + \delta = m$ gibt die Gesamtzahl der möglichen Fälle. Man findet nun $w(A, B) = \frac{\alpha}{m}$, $w(A) = \frac{\alpha + \beta}{m}$ und $w_A(B) = \frac{\alpha}{\alpha + \beta}$, denn die $\alpha + \beta$ Fälle, für die A sicher erfüllt ist, bleiben gleichmöglich, wenn A vorausgesetzt wird. Nun ergibt sich die Divisionsregel oder die Formel für die bedingte Wahrscheinlichkeit $w_A(B) = \frac{w(A, B)}{w(A)}$ für $w(A) \neq 0$ und damit die allgemeine Produktregel: $w(A, B) = w(A) w_A(B)$. Wenn A und B unabhängig sind, so ergibt sich mit $w_A(B) = w(B)$ die einfache Produktregel. Ist B von A unabhängig, so ist auch A von B unabhängig, denn beides führt zu $\alpha \delta - \beta \gamma = 0$.

Es folgt jetzt leicht auch die Umkehrung des letzten Axioms, d. h. die Formel $w_A(B) = w_A(C) \rightarrow w(A, B) = w(A, C)$ für $w(A) \neq 0$, denn es ist

$$w_A(B) = \frac{w(A, B)}{w(A)} \quad \text{und} \quad w_A(C) = \frac{w(A, C)}{w(A)}.$$

Um die Bedeutung des Aussonderungsaxioms noch besser klarzumachen, werde noch ein anderes Beispiel betrachtet, das auch an sich wichtig ist und das insbesondere zeigt, wie man nicht schließen darf; es ist das Beispiel der Bertrand'schen Kästchen.

Drei äußerlich gleiche Kästchen a, b, c enthalten je zwei Schubladen I und II. Im ersten Kästchen befindet sich in jeder Schublade ein Goldstück, im zweiten in der einen ein Gold-, in der andern ein Silberstück, im dritten je ein Silberstück. Die Kästchen werden gemischt, eines wird gezogen und eine Schublade geöffnet; man

¹⁾ A. A. MARKOFF, Wahrscheinlichkeitsrechnung. Leipzig und Berlin 1912. Seite 8.

findet darin ein Goldstück. Wie groß ist die Wahrscheinlichkeit, daß sich in der andern Schublade ein Silberstück befindet?

Die Voraussetzungen lauten $w(a) = w(b) = w(c)$ und $w_i(\text{I}) = w_i(\text{II})$ ($i = a, b, c$). Die Frage ist die nach der Wahrscheinlichkeit $w_O(b)$, wenn das Erscheinen eines Goldstückes mit O bezeichnet wird.

Hier darf man nun nicht so schließen: Die Kästchen sind ursprünglich gleichmöglich. Wenn ein Goldstück gefunden wurde, so sind nur noch a und b möglich, sie bleiben gleichmöglich, es wäre $w_O(b) = 1:2$. Dieser Schluß ist falsch, weil die Bedingung, daß ein Goldstück gefunden wird, nur für a sicher erfüllt ist, aber nicht für b ; hier hat man nur die Wahrscheinlichkeit $1:2$.

Der richtige Schluß geht so: Man zeigt zuerst, daß alle sechs Schubladen *a priori* gleichmöglich sind. Aus $w_a(\text{I}) = w_a(\text{II})$ folgt $w(a, \text{I}) = w(a, \text{II})$ nach der obigen Umkehrformel, außerdem ist $w(a) = w(a, \text{I oder } a, \text{II}) = w(a, \text{I}) + w(a, \text{II}) = 1:3$, also $w(a, \text{I}) = w(a, \text{II}) = 1:6$ und dasselbe gilt für b und c . Unter der Bedingung, daß ein Goldstück gefunden wurde, sind dann nur noch a, I , a, II und b, I möglich; diese bleiben aber gleichmöglich, weil für sie die Bedingung O sicher erfüllt ist. Es ergibt sich also $w_O(b) = 1:3$.

Für den Schulunterricht scheint mir eine Häufigkeitsrechnung für endliche Folgen als Einführung in die Wahrscheinlichkeitsrechnung ganz nützlich zu sein, denn es ergeben sich dabei Sätze und Beziehungen, die denen der Wahrscheinlichkeitsrechnung vollständig entsprechen. Das ist natürlich kein Zufall, es hat einen innern Grund. Man kann nämlich dem Begriff der Wahrscheinlichkeit eine genaue Häufigkeitsdeutung geben, nur nicht, wie es oft versucht wurde, in bezug auf eine zufällige, sondern in bezug auf eine bestimmte endliche Folge:

Die Wahrscheinlichkeit eines Ereignisses ist gleich der relativen Häufigkeit in einer Folge, in der alle gleichmöglichen Fälle gleich oft auftreten.

Die Häufigkeit in einer endlichen Folge ist unabhängig von der Anordnung; man braucht also keine Regellosigkeitsforderung, sondern kann sich die möglichen Fälle ganz regelmäßig angeordnet denken. Wenn man dann abzählt, wie viele davon für das Ereignis günstig sind, und dies durch die Gesamtzahl der Fälle dividiert, so hat man die klassische Definition $w = \frac{g}{m}$. Diese muß nur noch durch die Existenzforderung ergänzt werden, daß sich die möglichen Fälle tatsächlich in gleichmögliche zerlegen lassen und somit die zugehörigen Wahrscheinlichkeiten existieren, und dann noch durch die Forderung, daß gleichmögliche Ereignisse gleichmöglich bleiben, wenn eine für sie sicher erfüllte Bedingung vorausgesetzt wird. Damit ist die erwähnte Lücke ausgefüllt, und darauf läßt sich die ganze Wahrscheinlichkeitsrechnung aufbauen.

P. FINSLER, Zürich.

Kleine Mitteilungen

Ein Nomogramm zur normalen Axonometrie

In einer normalen Axonometrie gehe die Projektion der x -Achse senkrecht nach oben, die Projektionen der x - und y -Achse mögen mit der Waagrechten die Winkel σ und τ bilden. Die Bilder der Einheitsstrecken auf den drei Achsen sollen die Maßzahlen 1 , λ und μ haben. Es kann stets $\lambda, \mu \leq 1$ angenommen werden (Fig. 1).