Zufall und Mathematik: eine späte Liebe

von Ehrhard Behrends (FU Berlin)

Der Zufall ist sprichwörtlich „unberechenbar“. Wirklich hat es vergleichsweise lange gedauert, bis sich die Mathematiker des Themas angenommen haben.

Schnell wurde die Wichtigkeit des Gebiets erkannt, und entsprechend hoch war die Forschungsintensität. Jeder Überblick kann daher nur einen Teil der vielen interessanten Aspekte des Themas berühren. In der nachstehenden Auswahl wird die historische Entwicklung nur kurz gestreift. Unsere Themen:

- Wie fing es an?
- Wie macht man es heute?
- Grundlegende Konzepte
- Glücksspiel
- Der Zufall verliert sich im Unendlichen
- Die produktive Rolle des Zufalls
- Der Zufall im Mikrokosmos
- Philosophisches

1. Wie fing es an?

Vom Glücksspiel waren die Menschen schon immer fasziniert. Würfel finden sich in den ältesten Kulturen, und Tacitus berichtet von der besonderen Spielleidenschaft der Germanen.

Das Interesse der Mathematiker an den Gesetzen des Zufalls wurde aber erst im 17. Jahrhundert geweckt. Der Beginn der Wahrscheinlichkeitsrechnung wird üblicherweise auf den Zeitpunkt datiert, als Fermat gefragt wurde, wie man ein auf einen Sieg ausgesetztes Preisgeld bei vorzeitigem Abbruch eines Spiels aufteilen sollte.

Das ist überraschend schwierig, es lauern nämlich verschiedene Fallen bei der Berechnung der zugehörigen Wahrscheinlichkeiten. (Mit den gleichen Problemen haben auch heutige Studienanfänger zu kämpfen.)

Pierre de Fermat (1601 – 1665)

Ein Abriss der historischen Entwicklung ist hier nicht geplant. Nur soviel: Nach Fermat entwickelte sich das Gebiet recht schnell. Die grundlegenden Begriffe hatten sich bald herausgebildet, es gab die ersten wirklich tief liegenden Ergebnisse, und auch die ersten Standardwerke ließen nicht lange auf sich warten. Viele Mathematiker beschäftigten sich mit dem Gebiet, zu nennen sind insbesondere Jacob Bernoulli, Laplace, Gauß, ...

Trotzdem ist hervorzuheben, dass das Thema „Zufall" noch nicht wirklich in der Mathematik angekommen war. Viele Konzepte waren recht nebulös, es fehlte ganz einfach die Grundlage.

Es dauerte wesentlich länger als beim Thema „Zahlen" (da war das Fundament in der Mitte des 19. Jahrhunderts fertig), bis das überwunden war.

2. Wie macht man es heute?

Wie gehen die Mathematiker heute mit dem Zufall um, was verstehen sie unter „Wahrscheinlichkeit"? Als Illustration denken wir uns einen (unverfälschten) *Würfel*, den wir gleich werfen wollen. Welche Apekte dieser Situation sind typisch?

- Erstens ist klar, mit welchen Ergebnissen wir rechnen können: Zu erwarten ist eine der Zahlen von „1" bis „6".
- Welche beim Würfeln erscheinen wird, kann nicht vorausgesagt werden. Trotzdem zeigt die Erfahrung, dass man *bei häufiger Wiederholung* recht genaue Prognosen über die *Anteile* gemacht werden können: In etwa einem Sechstel der Fälle wird eine „3" zu erwarten sein, bei etwa der Hälfte der Würfe wird eine gerade Zahl erscheinen usw.
- Und mehr ist nicht zu erwarten: Auch wenn man tausend Mal mit einem Würfel gewürfelt hat, kann das Ergebnis beim Wurf Nummer 1001 nicht besser vorausgesagt werden als beim ersten Versuch.

Ganz ähnlich verhält es sich, wenn wir die *Lebensdauer einer Glühbirne* mathematitsch beschreiben wollen:

- Die möglichen Ergebnisse sind klar: irgendwelche positiven Zahlen (wir wollen vereinbaren, die Lebensdauer in Stunden zu messen).
- Wieder wird es bei vielen Versuchen „Tendenzen" geben: 80 Prozent der Glühbirnen hält länger als 100 Stunden, 10 Prozent gibt schon nach 20.4 Stunden den Geist auf usw.
- Und mehr kann nicht vorausgesagt werden: Niemand wird sagen können, wie viele Stunden eine frisch eingeschraubte Glühbirne brennen wird.

Es ist das Verdienst des russischen Mathematikers Andrey Kolmogoroff (1903 bis 1987),

die wesentlichen Aspekte des Zufalls, die auch in unseren beiden Beispielen betont wurden, extrahiert und als Axiomensystem herausgearbeitet zu haben. Das war 1933, dieses Jahr gilt als Beginn der modernen Wahrscheinlichkeitsrechnung.

Was ist nun ein *Wahrscheinlichkeitsraum* nach Kolmogoroff? Er besteht aus drei Dingen:

- Erstens aus der Menge Ω derjenigen Ergebnisse, die bei dem gerade zu beschreibenden Zufallsexperiment als Ergebnis zu erwarten sind[1]. Im Fall des Würfels besteht Ω aus der Menge $\{1, 2, 3, 4, 5, 6\}$ und im Glühbirnenbeispiel aus den positiven Zahlen. Ω kann aber auch viel einfacher sein (z.B. zweielementig, wenn als Ergebnis nur „Erfolg" oder „Misserfolg" zählt, etwa beim Versuch, mit einem Dart-Pfeil ins Schwarze zu treffen) oder viel komplizierter (wenn – wie bei der Beschreibung von Aktienkursen – Ω eine Menge von Funktionen ist.)

  Die Elemente aus Ω werden *Elementarereignisse* genannt.

- In fast allen Beispielen ist es eigentlich nicht wirklich wichtig, welches Elementarereignis sich denn nun gerade ergeben hat. Interessant kann vielmehr die Antwort auf die Frage „Wurde eine Zahl gewürfelt, die größer als Drei ist?" sein (beim Würfeln) oder „Wird die Lebensdauer mindestens 200 Stunden sein?" (bei den Glühbirnen). Deswegen führt man einen neuen, wichtigen Begriff ein: Ein *Ereignis* ist eine Teilmenge von Ω. Teilmengen werden also mit Fragen über den Ausgang des Versuchs identifiziert. (In den Beispielen traten die Ereignisse $\{4, 5, 6\}$ als Teilmenge von $\{1, 2, 3, 4, 5, 6\}$ und das Intervall $[\,200, \infty\,]$ als Teilmenge der positiven Zahlen auf.)

  Wir wollen hier der Einfachheit annehmen, dass *alle* Teilmengen von Ω als Ereignisse zugelassen sind[2].

- Nachdem wir nun erstens wissen, was Elementarereignisse sind und zweitens festgelegt haben, was wir unter Ereignissen verstehen wollen, kommt nun drittens der Wahrscheinlichkeitsbegriff hinzu. Jedem von uns betrachteten

[1] Der Buchstabe Ω – gesprochen „Omega" – hat sich als Bezeichnung für diese Menge eingebürgert.

[2] In Wirklichkeit ist es etwas komplizierter, denn es sind meist nur gewisse Teilmengen erlaubt. Das ist aus innermathematischen Gründen notwendig, weil man ausschließen muss, „zu viele" Ereignisse betrachten zu müssen. Für die Anwendungen ist das irrelevant, da alle sinnvollen Fragen immer gestellt werden können.

Ereignis E wird eine Zahl $P(E)$ (gesprochen „P von E") zugeordnet. Diese Zahl liegt zwischen Null und Eins und wird als Wahrscheinlichkeit dafür interpretiert, dass beim nächsten „Zufallsexperiment" das Ergebnis in E liegt.

Ist etwa $P(E) = 0.3$, so soll das bedeuten, dass in 30 Prozent der Experimente ein Elementarereignis erzeugt wird, das zu E gehört. (Will man eine Wahrscheinlichkeit in Mathematiker-Schreibweise in die im Alltagsgebrauch übliche Prozentsprache übersetzen, so muss man den Wert einfach mit 100 multiplizieren.)

Der Übergang von E zu $P(E)$ darf nicht völlig willkürlich sein, er soll ja eine Interpretation als Wahrscheinlichkeit erlauben. Man verlangt, dass $P(\Omega)$ gleich Eins ist (denn es passiert ja mit Sicherheit irgendetwas in Ω), und wenn sich Ereignisse E_1 und E_2 gegenseitig ausschließen, so soll die Wahrscheinlichkeit dafür, dass eines von beiden eintritt, die Summe aus den Einzelwahrscheinlichkeiten sein.

Das ist plausibel: Wenn im Glühbirnenbeispiel die Wahrscheinlichkeit für „Lebenszeit mindestens 200 Stunden" gleich 0.2 und für „Lebenszeit weniger als 20 Stunden gleich 0.1" ist, so sollte die Wahrscheinlichkeit für das Ereignis „Lebenszeit weniger als 20 Stunden oder mehr als 200 Stunden" gleich 0.3 sein.

In Kurzfassung besteht damit ein Wahrscheinlichkeitsraum aus einer Menge Ω, dem System der Ereignisse und einer Zuordnung mit „vernünftigen" Eigenschaften, die für jedes Ereignis E eine Zahl $P(E)$ zwischen 0 und 1 definiert.

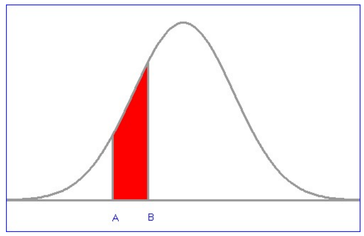
Zwei nahe liegende Fragen drängen sich auf. Erstens sieht das doch sehr trocken mathematisch aus, wo bleibt denn da das Mysteriös-Geheimnisvolle des Zufalls? Die Antwort ist für manche ernüchternd. Der Kolmogoroffsche Ansatz klammert nämlich alles Vage aus und beschränkt sich auf das mathematisch Beherrschbare. Die Situation ist damit vergleichbar mit dem Übergang von der aristotelischen Physik, die das „Wesen der Dinge" erforschen wollte, aber die wirklichen Probleme nur vor sich her schob, zur Beschränkung auf das im Experiment Messbare bei Galilei.

Und zweitens: Wie kann man denn die Zahlen $P(E)$ wirklich festlegen, es sind doch manchmal sehr, sehr viele Ereignisse zu berücksichtigen? In vielen Fällen ist das recht einfach, es gibt zwei große Klassen von Wahrscheinlichkeitsräumen, die den überwiegenden Teil der interessanten Situationen abdecken. Es könnte zum Beispiel sein, dass Ω *endlich* ist, etwa aus den Zahlen $1, 2, \ldots, 100$ besteht. Wenn man dann die Wahrscheinlichkeiten für die Elementarereignisse festgelegt hat (das wären 100 Definitionen), so hätte man damit auch die Wahrscheinlichkeiten für alle möglichen Ereignisse (das sind immerhin 2^{100} Mengen, eine gigantische Zahl[3]) definiert. Zum Beispiel muss die Wahrscheinlichkeit für das Ereignis $\{2, 4, 67\}$ die Summe aus den Wahrscheinlichkeitn für die Ereignisse $\{2\}$, $\{4\}$ und $\{67\}$ sein.

[3] 2^{100} ist immerhin gleich

$$1267650600228229401496703205376.$$

Und dann gibt es den Fall, dass Ω ein Intervall der Zahlengeraden ist und die Wahrscheinlichkeiten durch Angabe einer einzigen Funktion f erklärt werden können. Ist E ein Teilintervall (im Bild das Intervall zwischen den Punkten A und B), so soll $P(E)$ einfach die Fläche zwischen E und dem Graphen von f sein.



Zur Illustration haben wir das wichtige Beispiel der Normalverteilung gewählt, da ist f die so genannte *Glockenkurve*. (Die war – zu Ehren von Gauß – bis vor einigen Jahren auf jedem Zehnmarkschein abgebildet.) Man sieht, dass Ereignisse dann eine besonders hohe Wahrscheinlichkeit haben, wenn sie dort liegen, wo diese Kurve ihren Maximalwert hat.

3. Wichtige Konzepte

„Man sieht nur, was man weiß." So lautete das Motto des Baedeker-Reiseführers. Genauso ist es in der Mathematik. Es stellt sich oft erst nach vielen Jahrzehnten (oder gar Jahrhunderten) der Versuch-und-Irrtums-Phasen heraus, was die wirklich wichtigen Aspekte sind, auf die man besonders achten muss.

Hier einige Beispiele:

- *Bedingte Wahrscheinlichkeiten und Unabhängigkeit.* Mal angenommen, man wartet beim Würfeln auf eine Sechs, Die Wahrscheinlichkit dafür ist $1/6$. Wenn uns aber jemand – bevor das Ergebnis sichtbar ist – verrät, dass die gewürfelte Zahl größer als Drei ist, verändert sich unsere Erwartung. Es kommen jetzt nur noch die Zahlen $4, 5, 6$ in Frage, die Wahrscheinlichkeit für eine Sechs ist folglich auf ein Drittel gestiegen.

  Das ist ein spezieller Aspekt des Themas *Informationen verändern Wahrscheinlichkeiten.* Formalisiert sieht das so aus: Sind A und B Ereignisse, so soll $P(A \mid B)$ (gesprochen „Die Wahrscheinlichkeit von A unter der Bedingung B") die Wahrscheinlichkeit für A bezeichnen, wenn man schon weiß, dass B eingetreten ist. $P(A \mid B)$ hat den Wert

  $$\frac{P(A \cap B)}{P(B)},$$

  wobei $A \cap B$ das Ereignis „A und B treten beide ein" bezeichnet.

  Als Beispiel betrachten wir beim Ziehen aus einem Skatspiel die Ereignisse
  A: „Es wird ein König gezogen" und
  B: „Es wird ein Bild gezogen[4]".

[4] Wir wollen Bube, Dame und König als Bild zählen.

Es ist dann $P(A \cap B) = 1/8$, denn es gibt 4 Bilder, die Könige sind, in den 32 Karten. Klar ist auch, dass $P(B) = 3/8$, denn in einem Skatspiel ist der Anteil der Bilder $3/8$.

So folgt, dass $P(A \mid B) = P(A \cap B)/P(B) = (1/8)/(3/8) = 1/3$.

Zwei Ereignisse A und B heißen *unabhängig*, wenn $P(A \mid B) = P(A)$ gilt, wenn also die Information „Ergebnis in B“ die Wahrscheinlichkeit für A nicht verändert. Zum Beispiel sind, bei einem vollständigen Skatspiel, die Ereignisse „Es wird ein König gezogen“ und „Es wird eine Kreuz-Karte gezogen“ unabhängig: Sowohl $P(A)$ als auch $P(A \mid B)$ sind gleich $1/8$.

Es ist zu betonen, dass diese mathematischen Definitionen nur eine Formalisierung von Mechanismen sind, die wir im Alltagsleben täglich mehrfach antreffen können. Ein Beispiel für die Änderung von Wahrscheinlichkeiten durch Informationen: Sie lernen jemanden kennen und wollen wissen, ob er/sie an klassischer Musik interessiert ist. Der Anteil derartiger Personen in der Bevölkerung ist 20 Prozent.

Nun zeigt sich in einem der ersten Gespräche, dass er/sie Mozart für den Erfinder der Mozartkugeln hält. Das wird doch die Wahrscheinlichkeit, einen Klassikfan kennen gelernt zu haben, auf einen Wert nahe bei Null drücken.

Und allen ist klar, dass die Ereignisse „Er ist ein Klassikfan“ und „Er hat Schuhgröße 43“ unabhängig sind.

- *Zufallsvariable.* Eine Zufallsvariable ist so etwas wie eine *Informationskompression*. Ist etwa Ω die Menge aller möglichen Skatblätter, so könnte die Zufallsvariable angeben, wie viele Buben enthalten sind. Formal handelt es sich um eine Abbildung von Ω in irgendeine andere Menge, meistens wird – wie in unserem Beispiel – in den Bereich der Zahlen abgebildet.

  In diesem Fall kann man sich auch für den *Erwartungswert* interessieren. Der Erwartungswert einer Zufallsvariablen gibt an, auf welchen Wert man im Mittel rechnen sollte. Ein Beispiel: Wenn man mit Wahrscheinlichkeit 0.8 einen Gewinn von 10 Euro macht und mit Wahrscheinlichkeit 0.2 einen Verlust von 3 Euro, so wird man bei vielen Versuchen einen Mittelwert von $0.8 \cdot 10 - 0.2 \cdot 3$, also von 7.4 Euro erzielen.

  Der Erwartungswert ist als erste Maßzahl zur Beschreibung von Zufallsvariablen sehr nützlich, viele wesentliche Aspekte bleiben aber unberücksichtigt.

  Als Beispiel betrachten wir einen kleinen Jungen. Wenn er zu seinem Onkel geht, bekommt er immer 10 Euro geschenkt. Die Tante aber nimmt einen Würfel: Zeigt der eine „1“, bekommt der Junge 60 Euro, andernfalls gar nichts. In beiden Fällen ist der Erwartungswert 10 Euro, obwohl die Situationen sehr unterschiedlich sind.

- *Bedingte Erwartung.* Wie bei Wahrscheinlichkeiten gilt auch für Erwartungswerte, dass sie durch Informatonen verändert werden können. So ist zum Beispiel der Erwartungswert einer Würfelausgabe (mit gleicher Wahrscheinlichkeit wird eine der Zahlen $1, 2, 3, 4, 5, 6$ angezeigt) gleich 3.5. Bekomme ich aber die Information, dass die Ausgabe größer als 3 ist (dann sind nur noch die Zahlen $4, 5, 6$ mit gleicher Wahrscheinlichkeit zu erwarten), ist der Erwartungswert gleich 5.

- *Martingale.* Einer der wichtigsten Begriffe der modernen Wahrscheinlichkeitstheorie ist der Begriff des Martingals. Um das zu erklären, stellen wir uns eine Folge von Zufallsvariablen vor, jeweils eine für die Zeitpunkte $0, 1, 2, \ldots$ Ein typisches Beispiel wäre die Abfolge der Gewinne bei einem Glücksspiel, etwa beim Roulette. Bezeichnet X_n die n-te Zufallsvariable, so spricht man von einem *Martingal*, wenn die bedingte Erwartung von X_{n+1} jeweils gleich X_n ist. Das klingt sehr technisch, soll aber einfach bedeuten, dass das, was ich im nächsten Schritt zu erwarten habe, mehr oder weniger sein kann als das, was ich gerade habe, dass sich aber relative Gewinne und Verluste ausgleichen werden.

  Ein typisches Beispiel ist die Gewinnentwicklung bei einem fairen Spiel. Ist X_n der bis zur n-ten Runde erzielte Gewinn, so werden sich beim Übergang zum nächsten Spiel Gewinne und Verluste ausgleichen: Wenn ich es jetzt – im n-ten Spiel – auf 10 Euro gebracht habe, so werde ich auch eine Runde später *im Mittel* 10 Euro besitzen. Obwohl es im konkreten Einzelfall mehr oder weniger sein werden.

- *Stoppzeiten.* Bei diesem wichtigen Begriff ist die Grundidee einfach, die technische Umsetzung aber etwas technisch und daher für Neulinge in der Wahrscheinlichkeitsrechnung gewöhnungsbedürftig.

  Deswegen wird hier nur die Idee erläutert. Wir müssen uns eine Folge von Zufallsereignissen vorstellen, die nach und nach ablaufen: der jeweilige Gewinnstand in der n-ten Runde bei einer Folge von Glücksspielen, die Position eines Zufallsspaziergängers nach n Schritten, die Börsenkurse der Telekom-Aktie am n-ten Tag usw. Gleichzeitig häufen wir im Lauf der Zeit immer mehr Informationen an: Wir kennen die Gewinnentwicklung bis zum n-ten Spiel, den Weg des Zufallsspaziergängers bis zum n-ten Schritt, die Entwicklung der Telekom-Aktie seit Markteinführung. Und eine *Stoppzeit* ist dann nichts weiter als eine sinnvolle Vorschrift, zu irgendeinem Zeitpunkt „Stopp“ zu sagen. (Zum Beispiel, um dann das Casino zu verlassen oder die Aktie zu verkaufen.) Dabei soll „sinnvoll“ bedeuten, dass nur die bis zum potenziellen Stoppzeitpunkt angesammelten Informationen verwendet werden dürfen.

  Hier einige Beispiele für Stoppzeiten für das Glücksspiel-Beispiel:

  - Stoppe nach dem zehnten Spiel.
  - Stoppe drei Spiele, nachdem Du zum ersten Mal einen Gewinn von 100 Euro erzielt hast.
  - Stoppe, nachdem Du viermal einen Gewinn von mindestens 20 Euro hattest.

  Klar, dass man bei sorgfältiger Spiel-Beobachtung diese Stoppregel befolgen kann. *Keine* Stoppzeiten dagegen sind:

– Stoppe zwei Spiele, bevor der Gewinn ins Negative wechselt.
– Stoppe sofort, wenn für die nächsten 10 Spiele kein Gewinn von mindestens 100 Euro zu erwarten ist.

Das Traumziel aller Spieler wäre natürlich, eine Stoppzeit zu finden, die einem im Mittel einen Gewinn beschert. Geht das? Mehr dazu findet man im nächsten Abschnitt.

4. Glücksspiel

Wir wollen uns hier exemplarisch um zwei Glücksspiele kümmern, um Lotto und um Roulette. Zunächst zum Lotto.

Wenn der Jackpot hoch genug steht, stürmen die Medien die mathematischen Institute. Viele Ergebnisse kann man dabei schon recht elementar erhalten.

Es gibt 49 Felder, und deswegen hat man

$$49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44 = 9.620.865.408$$

Möglichkeiten, einen Lottoschein anzukreuzen.

Doch Halt! Die Reihenfolge spielt ja keine Rolle, und deswegen ist diese Zahl noch durch $6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$ zu teilen.

Insgesamt gibt es damit $13.983.816$ mögliche Lottoptipps.

Nur eine führt zum Hauptgewinn, und deswegen ist die Gewinnwahrscheinlichkeit die winzige Zahl $1/13.983.816$.

Wie kann man sich das vorstellen? Die Illustrationen sind leider etwas ernüchternd:

- *Die zufällige Telefonnummer:* Tippt man 7 Ziffern ganz zufällig ins Telefon, so gibt es dafür 10.000.000 Möglichkeiten. Wenn man hofft, eine ganz bestimmte zu wählen (z.B. die der netten Frau, die einem gerade im Bus gegenüber gesessen hat), so ist die Erfolgswahrscheinlichkeit $1/10.000.000$. Die ist noch deutlich höher als die Chance, einen Hauptgewinn im Lotto zu erzielen.
- *Der Stab an der Autobahn:* Stellen Sie sich eine Autobahnstrecke von 140 Kilometer Länge vor, das sind 14 Millionen Zentimeter. Wenn irgendwo ein Stab von einem Zentimeter Breite aufgestellt ist, dann ist die Chance, den im Vorbeifahren durch ein zufällig hinausgeschleudertes Centstück zu treffen, in etwa gleich der Lotto-Hautgewinnwahrscheinlichkeit.

Es ist leider hinzuzufügen, dass die Wahrscheinlichkeit für den Supergewinn noch durch 10 zu teilen ist! Jetzt müssten Sie auch bei einer achtstelligen Telefonnummer richtig liegen. Oder alternativ für zwei Personen auf Anhieb die vierstellige Bankautomaten-Geheimzahl erraten. Oder den Stab auf einer Strecke von 1400 Kilometer Länge (etwa Berlin-Rom) treffen.

Nun wollen wir ein Thema behandeln, das wesentlich anspruchsvoller ist: die *mathematische Theorie der Gewinnstrategien*. Schon in der ersten Hälfte des vorigen Jahrhunderts wurde streng bewiesen, dass man den Zufall nicht überlisten kann, doch hat sich das noch nicht allgemein herumgesprochen.

Mehr dazu gleich, doch es bietet sich an, hier zunächst eine dazu passende Geschichte aus dem Jahr 2005 zu erzählen:

- Im frühen Frühjahr ruft *stern-tv* an: Es wird ein Mathematik-Fachmann gebraucht, um eine Wette zu kommentieren. Herr G. hat behauptet, dass er beim Roulette immer gewinnen kann.
- Es gibt einige Drehtermine in der Spielbank Berlin und Livesendungen in Köln dazu.

Der mathematische Hintergrund: Es ist überhaupt keine Kunst, so zu spielen, dass man so gut wie immer mit einem kleinen Gewinn abschließt. Man darf allerdings nicht verschweigen, dass man Gefahr läuft, hin und wieder auch ganz bös zu verlieren.

Wir stellen uns ein faires Spiel vor: Man zahlt einen Euro Einsatz, dannn wird eine Münze geworfen; zeigt die „Kopf“, ist das Geld weg, und bei „Zahl“ bekommt man zwei Euro.

Mit der Strategie des Verdoppelns ist folgendes gemeint: Starte das Spiel mit einem Euro Einsatz. Im Fall des Gewinns verlasse die Spielbank (mit einem Euro Gewinn). Andernfalls setze zwei Euro, der Gesamteinsatz bisher ist also drei Euro. Im Fall eines Gewinns gibt es dann vier Euro, man kann also wieder mit einem Gewinn von einem Euro nach Hause gehen. Sollte es nicht geklappt haben, setzt man nun vier Euro usw. Der Einsatz wird also immer verdoppelt, bis einem das Glück einmal hold ist.

Der tröstliche Aspekt: Irgendwann *muss* ja wohl einmal „Zahl“ kommen. So ist etwa die Gefahr, zehnmal hintereinander zu verlieren, nur $1/2^{10} = 1/1024$, das ist weniger als ein Promille. Man kann also so gut wie immer mit einem Gewinn von einem Euro nach Hause gehen (wobei zwischendurch vielleicht mehrere hundert Euro eingesetzt wurden).

Allerdings gibt es auch den *beunruhigenden Aspekt:* Wenn man wirklich zehn Mal Pech hätte und es dann nicht weitergeht, wären 1023 Euro (der Gesamteinsatz beim Verdoppeln nach zehn Einsätzen) weg.

Dieses Risiko ist leider nicht zu vermeiden, denn irgendwann macht jede Spielbank zu, und außerdem werden nicht beliebig hohe Einsätze akzeptiert.

Im Fall von stern-tv hatte sich Herr G. eine Strategie zurechtgelegt, die einem versteckten Verdoppeln entsprach. Er hatte die Hoffnung – und die ging auch auf – dass ihm bei den etwa zehn Drehterminen das Glück immer hold sein würde. (Wirklich ist die Chance, dass es einmal nicht klappt, kleiner als ein Promille.)

Um ihn ein bisschen lächerlich zu machen, sollte gezeigt werden, dass das jeder kann. Dafür hat stern-tv wirklich keine Kosten gescheut. Es wurde extra ein Roulettetisch samt Croupiers aufgebaut, und ein Schimpanse sollte durch Zufallsentscheidungen so Roulette spielen, dass es dem Verdoppeln entsprach[5]. Fast erwartungsgemäß hat der Affe den Minimaleinsatz gewonnen ...

Die ganze Wahrheit ist etwas komplizierter. *Zunächst das Positive*: Man kann sich in weiten Grenzen bei einem fairen Spiel die Wahrscheinlichkeiten aussuchen, mit denen man gewinnen möchte. Man muss nur auf eine Balkenwaage, die nach links den Verlust und nach rechts den Gewinn anzeigt, 100 Gewichtseinheiten (die 100 Prozent entsprechen) so verteilen, dass die Waage im Gleichgewicht ist.

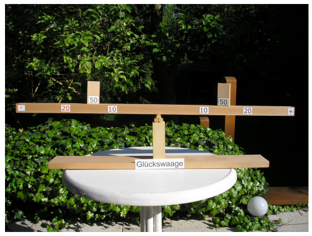
Mit geeigneten Strategien ist so eine Gewinn/Verlustverteilung dann zu realisieren. Hier einige Beispiele:

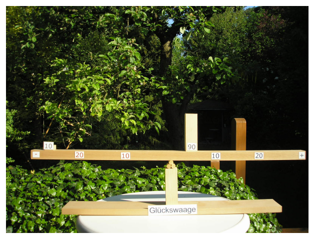
- Positioniere je 50 Prozent links und rechts in gleichem Abstand (z. B. 13 wie im Bild) vom Auflagepunkt. Das ist die Strategie der Vorsichtigen: Mit je 50 Prozent Wahrscheinlichkeit werden 13 Millionen Euro gewonnen und verloren.

- Lege 10 Prozent weit nach rechts (etwa bei 27) und 90 ein wenig nach links vom Auflagepunkt (etwa bei 3). Das ist die Standardsituation. Meist verliert man ein bisschen (3 Millionen Euro), aber hin und wieder gibt es den großen Wurf (27 Millionen Euro).

[5] Der Affe hat einfach gewürfelt, und je nach Ergebnis wurde auf „rouge“, „noir“, „manque“, „passe“, „pair“ oder „impair“ gesetzt.

- Und entscheidet man sich für 90 Prozent bei ein wenig rechts und 10 Prozent ziemlich weit links, so entspräche das der Strategie von Herrn G. bei stern-tv.





Und besser geht es auch nicht, mehr ist nicht zu erwarten. In der Sprache der Wahrscheinlichkeitstheorie liest sich das so: Wenn eine Folge von Zufallsvariablen ein Martingal ist (wenn also zum Beispiel die Gewinnentwicklung bei einem fairen Spiel beschrieben wird), so ist der Erwartungswert des Gewinns beim Stoppen gemäß irgendeiner realistischen Stoppzeit gleich dem Einstiegskapital[6]. Es ist also so, dass man das Spielen eigentlich gleich lassen könnte.

Die Moral: Der Zufall ist nicht zu überlisten, besser, man betritt die Spielbank erst gar nicht.

5. Der Zufall verliert sich im Unendlichen

Schon früh stellte man fest, dass sich Zufallseinflüsse gegenseitig aufheben können. Betrachten wir als Beispiel zwei Würfel. Es ist unwahrscheinlicher, eine 12 als Augensumme zu erhalten als eine 7, die Wahrscheinlichkeiten lauten $1/36$ gegen $1/6$.

Der Effekt wird noch wesentlich dramatischer, wenn viele Würfel beteiligt sind. Die Augensumme weicht dann nur mit einer minimalen Wahrscheinlichkeit bemerkenswert vom Mittelwert ab. (Bei n Würfeln hat er den Wert $n \cdot 3.5$.)

Es folgen einige quantitative Rechnungen. Wir werfen eine faire Münze zunächst 10 Mal. Es ist dann zu erwarten, dass „Kopf“ im Mittel fünf Mal vorkommt. Die genauen Werte lauten:

Die Kopfanzahl	Wahrscheinlichkeit dafür
ist exakt gleich 5	24.6%
liegt zwischen 4 und 6	54.2%
liegt zwischen 3 und 7	77.4%

[6] „Realistisch“ bedeutet dabei, dass das eingesetzte Kapital begrenzt ist und dass man nicht beliebig lange spielen kann.

Nun werfen wir 100 Mal:

Die Kopfanzahl	Wahrscheinlichkeit dafür
ist exakt gleich 50	7.95%
liegt zwischen 45 und 55	72.9%
liegt zwischen 40 und 60	96.5 %

Und im dritten Durchgang werfen wir 1000 Mal

Die Kopfanzahl	Wahrscheinlichkeit dafür
ist exakt gleich 500	2.52%
liegt zwischen 490 und 510	49.2%
liegt zwischen 480 und 520	80.6 %
liegt zwischen 470 und 530	94.6%

Wenn es also auch erwartungsgemäß recht unwahrscheinlich ist, dass es *genau* 500 Mal Kopf gibt, so kann man sich doch fast schon darauf verlassen (mit beinahe 95 Prozent Wahrscheinlichkeit), dass die Abweichung von dieser Zahl höchstens gleich 30 und damit der absolute Fehler kleiner als 6 Prozent ist.

Dieser Effekt wird bei der Überlagerung von noch mehr Zufallseinflüssen immer dramatischer. Die so genannten *Grenzwertsätze* beschreiben – auch quantitativ – dieses „Verschwinden des Zufalls". Solche Ergebnisse haben eine große praktische Bedeutung. Betrachten wir etwa das Problem eines Filialleiters in einem Supermarkt, am Sonnabend die richtige Menge Milch zu bestellen: Ist es zu wenig, sind die Kunden sauer, ist es zu viel, kann ein Teil der Ware am Montag nicht mehr verkauft werden. Deswegen wird man die Verkaufszahlen der Vergangenheit studieren, daraus die aktuellen Verkaufszahlen prognostizieren und dann so einkaufen, dass mit einer hohen Wahrscheinlichkeit alle Interessenten Milch kaufen können und der Verlust durch Restbestände klein ist.

Sinngemäß gelten die gleichen Überlegungen für die Planer der öffentlichen Verkehrsbetriebe, die entscheiden müssen, wie viele Wagen für den Zug um 7 Uhr 10 vorzusehen sind, und die Beipiele ließen sich natürlich beliebig vermehren.

Es folgen zwei weitere Beispiele, in denen die Grenzwertsätze eine wichtige Rolle spielen. Zunächst soll es um *Wahlprognosen* gehen. Wir stellen uns vor, dass die Bereitschaft eines Wählers, die XYZ-Partei zu wählen, durch eine noch unbekannte Wahrscheinlichkeit p gegeben ist, dass aber schon in einem Teil der Stimmbezirke die XYZ-Stimmen ausgezählt sind. Dann sollte aufgrund der Grenzwertsätze der Stimmanteil in der Gesamtbevölkerung recht genau gleich dem Anteil in diesen Stimmbezirken sein: So kommen die Voraussagen am Wahlabend zustande.

Etwas ausführlicher sollen *statistische Verfahren in der Qualitätskontrolle* beschrieben werden. Wir versetzen uns in die Rolle des Einkäufers einer Radiofabrik. Gerade eben sind 1000 Transistoren angekommen. Ist die Fehlerquote wirklich, wie vom Zulieferer zugesichert, kleiner als 3 Prozent?

Natürlich könnte man sie nun alle testen, aber erstens würde das viel zu lange dauern, und zweitens gehen Transistoren beim Testen manchmal kaputt. Deswegen werden nur 20 Exemplare einer eingehenden Prüfung unterzogen. Das Ergebnis: zwei sind defekt.

Der Einkäufer stellt nun die folgende Überlegung an. *Wenn* die Fehlerquote wirklich bei drei Prozent liegt, wie wahrscheinlich ist es dann, dass das, was beobachtet wurde, eintritt? Er erstellt eine kleine Tabelle, in der abzulesen ist, mit welcher Wahrscheinlichkeit $0, 1, \ldots$ defekte Stücke in einer Auswahl von 20 zu erwarten sind:

defekte Stücke:	0	1	2	3	4
Wahrscheinlichkeit dafür:	0.55	0.33	0.10	0.02	0.003

(Dass mehr als zwei defekte Stücke nur mit minimaler Wahrscheinlichkeit zu erwarten sind, ist eine Folgerung aus den Grenzwertsätzen.)

Man kann zum Beispiel ablesen, dass die Wahrscheinlichkeit für 2 defekte Stücke unter 20 zufällig ausgewählten gleich 0.1 (d.h. 10 Prozent) ist.

Das ist nicht besonders unwahrscheinlich, und deswegen gibt es keinen vernünftigen Grund, die Sendung zurückzuweisen. Ganz anders hätte es bei 4 defekten Transistoren ausgesehen. Das ist zwar auch nicht ausgeschlossen, denn unter den 1000 Transistoren dürfen sich ja 3 Prozent defekte, also 30 Stück verstecken, und durch Zufall könnten ja mehr als durchschnittlich zu erwarten in unsere Auswahl von 20 Testbeispielen gelangt sein. Das ist aber mit 3 Promille extrem unwahrscheinlich, und deswegen kann man nicht mehr so ohne weiteres von der Versicherung „Fehlerquote höchstens drei Prozent" ausgehen.

6. Die produktive Rolle des Zufalls

Ergebnisse von Zufallsexperimenten sind nicht vorhersagbar. Wie kann man da den Zufall produktiv einsetzen? Zu diesem Aspekt sollen einige Beispiele besprochen werden.

Wir beginnen mit dem *Comte de Buffon*, einem an der Wissenschaft interessierten Adligen, der im 18. Jahrhundert lebte.

Buffon war auf vielen Gebieten tätig, er brachte auch eine Enzyklopädie des damaligen Wissens heraus. Hier geht es aber um ein berühmtes Experiment aus der Wahrscheinlichkeitstheorie.

Man braucht dazu einen Dielenfußboden und ein nicht zu langes Stöckchen. Dann kann man sich fragen:

> Mit welcher Wahrscheinlichkeit fällt das Stöckchen so, dass eine Dielenkante gekreuzt wird?

Das kann man mit etwas Elementargeometrie ausrechnen, wenn man den Zufall richtig modelliert (Mittelpunkt des Stöckchens und Drehwinkel fallen völlig gleichmäßig). Es ergibt sich: Die Wahrscheinlichkeit P ist

$$P = \frac{2 \cdot l}{\pi \cdot d},$$

dabei ist π die Kreiszahl, l die Stöckchenlänge und d die Dielenbreite. (Die Zahl π taucht hier deswegen auf, weil die Drehwinkel – vom Mittelpunkt eines Kreises mit Radius Eins aus gesehen – in die Richtung aller Punkte auf der Kreisperipherie zeigen können. Und diese Peripherie hat die Länge 2π.)

Buffon hatte nun die bemerkenswerte Idee, das nach π aufzulösen:

$$\pi = \frac{2 \cdot l}{P \cdot d}.$$

Damit ergibt sich die Möglichkeit, zunächst die Wahrscheinlichkeit P approximativ durch fleißiges Stöckchenwerfen zu bestimmen und daraus einen Näherungswert für π herzuleiten.

Damit war das erste *Monte-Carlo-Verfahren* der Mathematikgeschichte erfunden. Darunter versteht man Algorithmen, bei denen der Zufall zu Berechnungen eingesetzt wird. Inzwischen gehören solche Verfahren zum Standardrüstzeug. Sie sind bequem zu programmieren und dank der heutigen Computermöglichkeiten auch sehr effektiv.

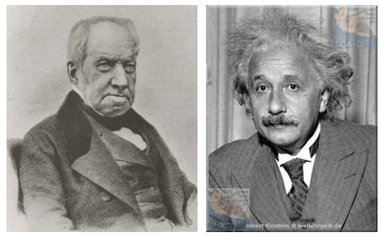
Als weiteres Beispiel von Monte-Carlo-Verfahren betrachten wir *Flächenberechnungen*: Um die Fäche einer irgendwie in einem Rechteck liegendenn Figur zu bestimmen, muss man nur „sehr viele“ Zufallszahlen in diesem Rechteck erzeugen und den prozentualen Anteil der Treffer (in der Figur) protokollieren. Hat das Rechteck etwa die Fläche 1 und wurden bei 10000 Versuchen 6534 Treffer gezählt, so sollte die Figur etwa die Fläche 0.65 haben.

Oft ist es so, dass andere Verfahren nicht zur Verfügung stehen. Als gravierender Nachteil bleibt aber immer, dass die Ergebnisse nie sicher sind! Eine typische Aussage – etwa im Zusammenhang mit dem Buffon-Experiment – lautet: „Mit 98 Prozent Wahrscheinlichkeit ist die Zahl π bis auf drei gültige Ziffern durch 3.14 gegeben.“ Es ist offensichtlich, dass man mit solchen Aussagen nicht viel anfangen kann, wenn man ein Ergebnis wirklich genau wissen muss.

> In dem Betrag über Quantencomputer des gleichen Autors in diesem Buch geht es übrigens auch darum, Zufallsverfahren zur Lösung von Problemen einzusetzen. Die „Erfolgsquote“ ist da recht bescheiden (50 Prozent), doch ist das bei der Entschlüsselung von Geheimcodes kein gravierender Nachteil. Bei 10 Versuchen müsste man schon sehr viel Pech haben, um den Code *nicht* zu knacken. (Die Wahrscheinlichkeit für so viel Pech ist nur $1/2^{10}$, also etwa ein Promille.)

7. Der Zufall im Mikrokosmos

Wir betrachten zunächst die *Brownsche Bewegung*. Der Botaniker Robert Brown hatte 1827 beobachtet, dass Blütenpollen unter dem Mikroskop eine Irrfahrt durchführen. Für viele Jahrzehnte wurde das Phänomen nicht verstanden, bei den Erklärungsversuchen gab es wilde Spekulationen (wie etwa die Annahme einer „Lebenskraft“ in der anorganischen Materie).



Brown und Einstein

Um 1900 gab es dann schon einen ersten Versuch, die Brownsche Bewegung zur Modellbildung heranzuziehen. Der Franzose Bachelier beschrieb damit die Entwicklung von Aktienkursen. (Er war damit seiner Zeit weit voraus, die Bedeutung dieser Modelle für die Finanzmathematik wurde erst vor wenigen Jahrzehnten erkannt.)

Viel wichtiger für die Entwicklung der Wissenschaft war aber eine Arbeit von Albert Einstein zur Brownschen Bewegung aus dem Jahr 1905. Darin hatte Einstein einmal durchgerechnet, wie sich ein mikroskopisch kleines Teilchen bewegen würde, wenn es von Molekülen zufällig angestoßen werden würde. Aus dieser Arbeit waren auch quantitative Voraussagen ableitbar, die in den darauffolgenden Jahren bestätigt wurden. (Nach n Zeiteinheiten sollte sich das Teilchen – bis auf einen Faktor – im Mittel um $\sqrt{n}$ Längeneinheiten vom Startpunkt entfernt haben.)

Damit wurde der damals noch offene Streit entschieden, ob die Welt im Kleinen nun aus Teilchen aufgebaut ist oder aus einem kontinuierlichen Medium besteht: Die Annahme kleinster Teilchen führt zu makroskopisch beobachtbaren Phänomen, die durch Eigenschaften eines Kontinuums nicht erklärbar sind, und das erlaubt es, von zwei konkurrierenden Theorien eine zu verwerfen.

Heute ist die Brownsche Bewegung in der Wahrscheinlichkeitstheorie allgegenwärtig. Bemerkenswerter Weise ist die präzise Behandlung recht aufwändig, Mathematikstudenten lernen sie erst im zweiten Semester ihrer Beschäftigung mit der Wahrscheinlichkeitstheorie kennen. Es dauerte auch bis in die zwanziger Jahre des vorigen Jahrhunderts, bis – durch den amerikanischen Mathematiker Norbert Wiener – gezeigt werden konnte, dass die verschiedenen Anforderungen an ein mathematisches Modell der Brownschen Bewegung gleichzeitig verwirklicht werden können (stetige Pfade, Unabhängigkeit der Zuwächse, linear wachsende Varianz, normalverteilt im Raum).

Für Mathematiker ist dieser Punkt – Nachweis der Existenz eines Modells mit den gewünschten Eigenschaften – fundamental wichtig und leider ziemlich schwierig. Für Anwender, etwa aus dem Ingenieurbereich, ist die Realität der Brownschen Bewegung sonnenklar, da sie in der täglichen Arbeit allgegenwärtig ist (etwa als Rauschen bei der Signalübertragung).

Dieses unterschiedliche Bedürfnis nach theoretischer Absicherung (Mathematiker gegen Ingenieure) zeigte sich in einer für den Autor überraschenden Deutlichkeit auf einer Tagung zur Signalverarbeitung (Maspalomas 2006), als während eines historischen Vortrags die Bemühungen der Mathematiker um Strenge von der Mehrheit des Auditoriums nur mit ungläubigem Staunen zur Kenntnis genommen wurden.

Heute ist die Brownsche Bewegung auch in der Finanzmathematik allgegenwärtig. Durch die Verleihung des Nobelpreises für Wirtschaft im Jahr 1997 an Black, Scholes und Merton wurde nämlich einer größeren Wissenschaftlergemeinde klar, dass man damit hervorragend die Entwicklung von Aktienkursen modellieren kann. Dass hatte zwar Bachelier auch schon um 1900 versucht, aber erst viele Jahrzehnte später war dieses Modell bei der Bewertung von Optionen unverzichtbar[7].

Der zweite große Aspekt des Themas ist die *Quantenmechanik*. Etwas vereinfacht kann man sich die heute am weitesten verbreitete Interpretation so vorstellen, dass die Mikrowelt von Wahrscheinlichkeiten regiert wird.

Wenn man keine Messungen vornimmt, entwickeln sich diese Wahrscheinlichkeiten auf genau berechenbare Weise. Sobald aber gemessen wird (Energie, Impuls, durch welchen Spalt, gespiegelt oder nicht?), ist es so, als ob ein Zufallsgenerator für das Ergebnis verantwortlich wäre. Bemerkenswert ist dabei, dass die Wahrscheinlichkeiten durch Pfeile in der Ebene verschlüsselt sind[8]: Die relevante Wahrscheinlichkeit ist das Quadrat der Pfeillänge. Und da die Überlagerung zweier Situationen durch die Addition der Pfeile gemäß der Vektoraddition beschrieben wird, kann es passieren, dass – wenn die Pfeile in entgegengesetzte Richtung zeigen – sich bei Überlagerung aus zwei großen Wahrscheinlickeiten sehr kleine Wahrscheinlichkeiten ergeben.

8. Philosophisches

Manche Leser werden sich gewundert haben, dass die Frage „Was ist Wahrscheinlichkeit?“ ausgeklammert wurde. Weiter oben wurde schon betont, dass sich dieser Verzicht als sehr erfolgreich herausgestellt hat. Das ist ganz ähnlich wie – zum Beispiel – in der Geometrie. Dort werden ja auch die Begriffe „Punkt“, „Gerade“ usw. ohne nähere Erklärung an den Anfang gestellt, es gelten gewisse Tatsachen („Zwischen je zwei verschiedenen Punkten gibt es genau eine Gerade“ usw.) und dann kann es losgehen.

Das mathematisches Modell zur Behandlung des Zufalls (die Wahrscheinlichkeitsräume) hat sich in dem Sinne als angemessen erwiesen, als die sich daraus ergebenden Voraussagen mit dem, was in der von uns so genannten Wirklichkeit

[7] Mehr dazu findet man im Artikel von Walter Schachermayer in diesem Buch.
[8] *So* wollen wir uns komplexe Zaheln vorstellen.

passiert, recht gut übereinstimmen. Versicherungsunternehmen, die sich nach den Gesetzen der Wahrscheinlichkeitstheorie richten, können sich auf die vorausgesagten Zufalls-Schwankungen verlassen, der Staat kann sicher sein, dass die Lizenz des Spielcasinos nicht zu exorbitanten Gewinnen der Betreiber führt usw.

Trotzdem hat es natürlich Versuche gegeben, Wahrscheinlichkeit und Zufall etwas genauer zu fassen. Als Beispiel soll hier auf den Häufigkeitsansatz von Richard von Mises (1883-1953) hingewiesen werden. Nach von Mises muss man, um zur Wahrscheinlichkeit eines Ereignisses E zu gelangen, sehr oft testen, ob bei dem gerade relevanten Zufallsexperiment das Elementarereignis in E zu finden ist. Ist das bei n Versuchen h_n Mal der Fall, so soll die Wahrscheinlichkeit von E als der Grenzwert $\lim_{n\to\infty} h_n/n$ definiert sein.

Leider kann niemand garantieren, dass dieser Grenzwert existiert, und selbst wenn, ist nicht klar, dass bei verschiedenen Versuchsreihen das gleiche Ergebnis herauskommt.

In diesem Abschnitt soll auch noch auf ein anderes Problem eingegangen werden, das immer wieder für Verwirrung sorgt. Man stelle sich einen Würfel vor, der ganz oft geworfen wird. Einerseits wird immer gesagt, dass bei „vielen“ Versuchen im Mittel alle Zahlen etwa gleich oft vorkommen. Gleichzeitig wird aber auch behauptet, dass der Zufall kein Gedächtnis hat, die Chancen also bei jedem neuen Wurf genauso sind wie am Anfang.

Das kann doch irgendwie nicht stimmen: Wenn man sehr oft gewürfelt hat und keine Sechs dabei war, muss sich der Würfel doch wohl ein bisschen anstrengen und verstärkt Sechsen produzieren, um die erste Forderung zu erfüllen; die Chancen für „Sechs“ sollten also deutlich steigen. Entsprechend setzen viele beim Lotto ja auch auf diejenigen Zahlen, die lange nicht gezogen wurden.

Der Widerspruch löst sich dadurch auf, dass die Chancengleichheit aller Zahlen genau genommen kein Muss ist, sondern nur mit überwältigender Wahrscheinlichkeit erwartet werden darf. Man kann berechnen, dass bei einem Würfelexperiment die Chancen nahe bei 100 Prozent sind, dass alle Zahlen in etwa gleich oft auftreten. Es ist aber – mit unglaublich kleiner Wahrscheinlichkeit – durchaus möglich, dass etwas Unerwartetes passiert, dass etwa nur Dreien gewürfelt werden.

Zum Abschluss gibt es hier noch einige Literaturempfehlungen:

* Jörg Bewersdorff: Glück, Logik und Bluff (Vieweg Verlag 2001).
In diesem Buch wird das Thema „Glücksspiele“ vertieft.

* Hans-Otto Georgii: Stochastik (de Gruyter Lehrbuch 2002).
Dieses Buch kann allen empfohlen werden, die sich ernsthafter mit Stochastik auseinandersetzen wollen. Es ist für Studierende aller Fachrichtungen und interessierte Schüler der Oberstufe geeignet.

* Ellen Kaplan - Michael Kaplan: Eins zu Tausend (Campus Verlag 2007).
Hier finden alle, die mehr über die historische Entwicklung und die Beziehungen zur Statistik erfahren wollen, eine überwältigende Fülle von interessanten Informationen.

* Rudolf Taschner: Zahl, Zeit, Zufall (Ecowin Verlag 2007).
Der Autor fasst die mathematische Theorie des Zufalls als Teil der Beschreibung der Welt durch Mathematik auf. Auch hier gibt es viele historische Hintergrundinformationen, die Mathematik beim Roulette wird auch ausführlich behandelt.

Ehrhard Behrends
Fachbereich Mathematik und Informatik der FU Berlin
e-mail: behrends@math.fu-berlin.de