

Schüler-SimuLab
Kursreihe stochastische
Simulationen
Kurs 3

Erste stochastische Simulationen

Stefan Hartmann
Forschungszentrum caesar

18. Januar 2008

Ein paar einleitende Worte

Im ersten Kurs haben wir uns mit der Erzeugung von Pseudozufallszahlen beschäftigt. Dort haben wir auch schon eine erste kleinere Anwendung kennengelernt, nämlich die Simulation eines Würfels. Man musste dabei eine Menge von Zufallszahlen aus dem Intervall $[0, 1[$ auf die Menge $\{1, 2, 3, 4, 5, 6\}$ transformieren. In diesem Kurs wollen wir weitere Anwendungen der stochastischen Simulation kennenlernen, die schon etwas anspruchsvoller sind.

Wir werden sehen, dass man stochastische Simulationen dazu nutzen kann, um stochastische Berechnungen vorzubereiten (d.h. mit ihrer Hilfe Ideen für einen Ansatz zu finden), bereits vollzogene Berechnungen zu plausibilisieren oder aber diese sogar zu ersetzen, wenn der direkte mathematische Weg zu schwierig oder aber sogar unmöglich ist.

Als Beispiele werden wir mit Hilfe des sogenannten Galton-Bretts die Binomialverteilung stochastisch simulieren, Flächeninhalte simulativ berechnen (und dabei mit der Schätzung von π die wohl „klassischste“ und bekannteste aller Monte-Carlo-Simulationen kennenlernen) sowie Erwartungswerte simulativ bestimmen. Anschließend werden wir mit Hilfe stochastischer Simulationen zwei auf den ersten Blick unintuitive stochastische Phänomene plausibilisieren. Im ersten Fall handelt es sich um das sogenannte „Würfelschlängenphänomen“, das man im „Mathematikum“, einem Mathematikmuseum in Gießen mit Herrn Professor Dr. Beutelspacher als Leiter, sowie im Rahmen der Wanderausstellung „Mathematik zum Anfassen“ als echtes Experiment durchführen kann. Das zweite hier vorgestellte Beispiel zählt wohl zu den bekanntesten Problemen überhaupt: Es handelt sich um das sogenannte „Ziegenproblem“: Man betrachtet ein Quizspiel, bei dem hinter zwei Türen eine Ziege und hinter der dritten ein Auto steckt. Der Kandidat entscheidet sich für eine Tür, woraufhin der Quizmaster (der natürlich weiß, wo sich das Auto befindet) eine Tür öffnet, hinter der sich eine Ziege verbirgt. Ist es nun besser für den Kandidaten bei seiner ersten Wahl zu bleiben oder aber die Tür zu wechseln? Diese Frage werden wir theoretisch beantworten,

aber - da auch nach Kenntnis der Begründung meist Restzweifel bleiben - auch simulativ verifizieren.

Ich denke, dass ihr bei allen Beispielen seht, welchen Sinn und Zweck stochastische Simulationen haben können und welch wichtiges und mächtiges Werkzeug sie bei Überlegungen im Umfeld der Stochastik darstellen. Weitere Anwendungen werden wir in weiteren Kursen kennenlernen. Ich denke da in erster Linie an Ruinwahrscheinlichkeiten und ähnliches sowie Simulationen im Rahmen der Finanzmathematik (etwas Simulationen von Aktien- und Zinskurven). Die relativ überschaubaren, aber zum Teil schon recht anspruchsvollen Anwendungen in diesem Kurs sollen einen ersten Vorgeschmack auf komplexere Anwendungen geben.

In jedem Fall bin ich davon überzeugt (zumindestens habe ich große Hoffnung), dass euch der Kurs viel Spaß bereitet und ihr einen ersten Eindruck davon bekommt, an welchen Stellen man die Mathematik (und speziell die stochastische Simulation) auch bei praktischen Überlegungen effektiv einsetzen kann.

Bonn, den 20.12.2004

Stefan Hartmann

Kapitel 1

Was ist eine stochastische Simulation?

Eine präzise Definition dieses weitläufigen Begriffs ist nicht einfach. Sehr allgemein lassen sich aber stochastische Simulationen wie folgt beschreiben:

Definition 1 (stochastische Simulation)

*Eine **stochastische Simulation** ist eine Nachbildung eines stochastischen Systems auf einem Computer zwecks Untersuchung der Eigenschaften dieses Systems.*

Damit unterscheidet sich also eine stochastischen Simulation sowohl von einem **echten empirischen (d.h. aus Beobachtungen und Erfahrungen gewonnenen) Experiment**, bei dem man einen Vorgang in der Natur oder der Wirtschaft beobachtet und Daten erhebt, als auch von einer **rein mathematisch-analytischen Untersuchung**.

*empirisch:
von griech.
émpeiros =
erfahren*

Gemeinsam mit einem echten empirischen Experiment ist **der empirische Ansatz** (d.h. das Zählen, Messen, etc.), gemeinsam mit einer rein mathematischen Untersuchung ist ein zu Grunde liegendes **mathematisches Modell** (als Abbild der Wirklichkeit).

Beispiel:

Wir betrachten die **Lagerhaltung** einer Firma mit verschiedenen Produkten, d.h. wir fragen uns, welche Menge von jedem einzelnen Produkt sich zu gewissen Zeitpunkten im Lager befindet. Die Lagerhaltung hängt einerseits ab von der **Nachfrage** (Zu welchen Zeitpunkten wird ein Produkt in welcher Menge verlangt?) und andererseits von der **Lagerhaltungspolitik** (Wie groß ist das Lager und wann wird nachbestellt?) sowie den **Lieferfristen** (Wie lange dauert es von der Bestellung zur Auslieferung ans Lager?). Nachfrage und Lieferfristen werden meist stochastisch modelliert.

Das Ziel ist die Optimierung der Lagerhaltungspolitik im Hinblick auf die Kosten und auf die Wahrscheinlichkeit, dass ein verlangter Artikel im Lager ist. Typischerweise stellt sich nach einer „chaotischen“ Anfangsphase ein Gleichgewicht ein. Bei einer rein mathematisch-analytischen Untersuchung würde man versuchen, die Wahrscheinlichkeit, dass ein verlangter Artikel zu gewissen Zeitpunkten im Lager ist sowie die Wahrscheinlichkeiten für verschiedene Lagerungsdauern explizit zu berechnen. Bei einer stochastischen Simulation erzeugt man zufällige Nachfragezeitpunkte und -mengen nach den einzelnen Produkten und zählt dann, wie oft ein verlangter Artikel nicht im Lager ist.

Wieso macht man überhaupt Simulationen?

- Der Aufwand (an Zeit und Geld) ist meist kleiner als bei einem echten Experiment (zum Beispiel kann man in dem obigen Beispiel sehr leicht die Lagerhaltungspolitik ändern, was man in der Realität nur mit großem Aufwand tun kann).
- Die Komplexität eines Systems ist meist größer als bei einer rein mathematisch-analytischen Untersuchung, d.h. man kann Szenarien untersuchen, die man auf rein mathematisch-theoretischem Wege nicht in den Griff kriegen würde, weil dann die Formeln sehr kompliziert würden und man gewisse mathematische Terme nur mühsam (oder gar nicht) berechnen könnte.

Wir wollen im Folgenden speziell stochastische Simulationen dafür nutzen, Wahrscheinlichkeiten von gewissen Ereignissen und auch Flächeninhalte ap-

proximativ (=näherungsweise) zu berechnen, deren Berechnung auf rein mathematisch-analytischem Wege teilweise sehr aufwändig wäre.

Zur Erläuterung: Flächeninhalte kann man auch als Wahrscheinlichkeiten interpretieren, wie wir noch genauer sehen werden. In der Schule nennt man solche Wahrscheinlichkeiten häufig **geometrische Wahrscheinlichkeiten**.

Auf diese Weise kann man umgekehrt auch Ergebnisse, die man auf direktem Wege errechnet hat, mit Hilfe von Simulationen noch einmal überprüfen.

Die **Idee bei einer stochastischen Simulation** soll nun in allgemeiner Form ausführlich erläutert werden:

Wir haben zunächst einmal ein theoretisches Experiment, dessen Ausgang vom Zufall abhängt und das man mathematisch wie folgt modellieren kann:

Unter einem **Elementarereignis** ω versteht man den möglichen Ausgang einer Messung, Beobachtung usw. Die Gesamtheit der Elementarereignisse bildet den **Ereignisraum** Ω (oder die **Ereignismenge** Ω). Wir gehen zunächst einmal davon aus, dass Ω nur aus endlich vielen Elementarereignissen besteht, denn das vereinfacht die folgenden Betrachtungen erheblich. In diesem Fall werden **alle** Teilmengen A von Ω , die im Allgemeinen mehr als ein Elementarereignis enthalten können, als **Ereignisse** bezeichnet.

Beispiel:

Beim Würfeln mit einem Würfel entsprechen die sechs möglichen Elementarereignisse dem Auftreten einer bestimmten Augenzahl, während z.B. das Ereignis „gerade Augenzahl“ die Tatsache bezeichnet, 2, 4 oder 6 Augen geworfen zu haben. Hier ist also $\Omega = \{1, 2, 3, 4, 5, 6\}$ der Ereignisraum und $A = \{2, 4, 6\}$ beschreibt das Ereignis „man würfelt eine gerade Augenzahl“.

Spezielle Ereignisse sind immer die leere Menge $\emptyset$ (das **unmögliche Ereignis**) und der ganze Ereignisraum Ω (das **sichere Ereignis**).

Man kann nun eine Abbildung P definieren, die jedem Ereignis A einen Wert $P(A)$ zuordnet, mit dem die Wahrscheinlichkeit modelliert wird, dass das Ereignis A eintritt. Dabei müssen die folgenden Bedingungen erfüllt sein:

1. $0 \leq P(A) \leq 1$ für alle Ereignisse A ,
2. $P(\emptyset) = 0, P(\Omega) = 1$,
3. $P(A \cup B) = P(A) + P(B)$, wenn $A \cap B = \emptyset$,

also Forderungen, die man von einer Wahrscheinlichkeit auch erwartet!

Hierbei ist für zwei Mengen $A, B \subset \Omega$ die Menge $A \cap B$ der **Durchschnitt** von A und B (also die $\omega \in \Omega$, die sowohl in A als auch B enthalten sind) und die Menge $A \cup B$ die **Vereinigung** von A und B (also die $\omega \in \Omega$, die in A oder B (oder beiden) enthalten sind). Man nennt zwei Mengen disjunkt, wenn es keine Elemente gibt, die in beiden enthalten sind, d.h. wenn die beiden Mengen einen leeren Durchschnitt haben, wenn also der Fall $A \cap B = \emptyset$ eintritt. Die dritte Bedingung bedeutet: Wenn man wissen will, wie groß die Wahrscheinlichkeit dafür ist, dass das Ereignis A oder das Ereignis B eintritt, dann kann man die beiden Wahrscheinlichkeiten für A und B einfach addieren, wenn man genau weiß, dass nicht A und B zugleich eintreten können. Sobald es ein $\omega \in \Omega$ gibt mit $\omega \in A \cap B$, geht das nicht mehr. Dort wird die Formel etwas komplizierter:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

wie man sich zumindestens anschaulich leicht klarmacht (die Fälle, wo A und B zugleich vorkommen, werden in $P(A) + P(B)$ doppelt gezählt und müssen daher anschließend einmal angezogen werden).

(Weiterführende Bemerkung: Bei unendlichen Mengen wird die dritte Bedingung erweitert. Dort betrachtet man eine ganze (unendliche) Folge paarweise disjunkter Ereignisse, also eine Folge $(A_n)_{n \in \mathbb{N}}$ von Ereignissen A_n mit $A_i \cap A_j = \emptyset$ für alle $i, j \in \mathbb{N}$ mit $i \neq j$ und verlangt dann $P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} P(A_n)$.)

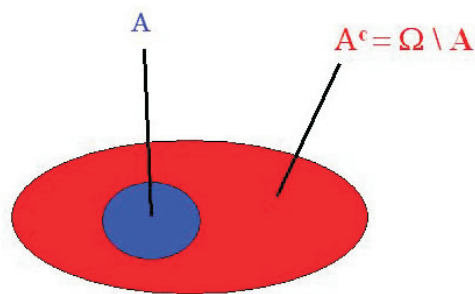
Man bezeichnet im Allgemeinen mit $\mathcal{P}(\Omega)$ die sogenannte **Potenzmenge**

von Ω , also die Menge, die alle Teilmengen von Ω enthält. Ist Ω (wie hier) endlich, dann besteht $\mathcal{P}(\Omega)$ also aus allen möglichen Ereignissen. In diesem Fall sagt man, dass eine Funktion $P : \mathcal{P}(\Omega) \rightarrow [0, 1]$ mit den obigen drei Eigenschaften ein **Wahrscheinlichkeitsmaß** ist.

Gruppenaufgabe:

Leite aus diesen Folgerungen die Aussage her, dass für das **Gegenereignis** $A^c := \Omega \setminus A := \{\omega \in \Omega : \omega \notin A\}$ eines Ereignisses A die folgende Aussage gilt:

$$P(A^c) = 1 - P(A).$$



Gegenereignis als Komplement einer Menge

Also: Wenn man die Wahrscheinlichkeit dafür berechnen will, dass ein Ereignis A nicht eintritt, kann man einfach die Wahrscheinlichkeit dafür berechnen, dass A eintritt und diesen Wert von 1 abziehen. Ist $A \subset \Omega$ ein Ereignis, so nennt man, wie in der obigen Aufgabe, $A^c = \Omega \setminus A$ das **Gegenereignis** zu A . In vielen Fällen ist es einfacher über das Gegenereignis zu rechnen und dann diesen Wert von 1 abzuziehen.

Beispiel:

Wie groß ist die Wahrscheinlichkeit mit zwei Würfeln als Augensumme eine größere Zahl als 2 zu würfeln?

Misst X die Augensumme, dann könnte man so direkt rechnen:

$$P(X > 2) = P(X = 3) + P(X = 4) + \dots + P(X = 12).$$

In diesem Fall müsste man die einzelnen Wahrscheinlichkeiten alle ausrechnen und dann addieren.

Schneller ginge es über das Gegenereignis (man beachte dabei, dass die Augensumme nicht kleiner als 2 werden kann und daher $P(X \leq 2) = P(X = 2)$ gilt):

$$P(X > 2) = 1 - P(X \leq 2) = 1 - P(X = 2) = 1 - \frac{1}{36} = \frac{35}{36}.$$

Wenn Ω endlich ist, dann genügt die Angabe von $P(\{\omega\})$ für alle Elementarereignisse ω , um die Wahrscheinlichkeiten $P(A)$ für alle Ereignisse A zu berechnen, denn für $A = \{\omega_1, \omega_2, \dots, \omega_n\} \subset \Omega$ gilt ja:

$$P(A) = P(\{\omega_1\}) + P(\{\omega_2\}) + \dots + P(\{\omega_n\}),$$

indem man sukzessive die Eigenschaft 3 anwendet.

Wir wollen mit Hilfe einer stochastischen Simulation näherungsweise die Wahrscheinlichkeit dafür bestimmen, dass A eintritt. Dazu führen wir nun das theoretische Experiment immer wieder durch, und zwar nicht wirklich im Sinne eines echten empirischen Experimentes, sondern mit Hilfe einer stochastischen Simulation, also mit Hilfe von Zufallszahlen. Bei jedem Simulationsschritt verwendet man andere vorher generierte Zufallszahlen, so dass es jedesmal zu unterschiedlichen Szenarien (möglichen Ausgängen des Experimentes) kommt. Wichtig ist dabei natürlich, dass man die Zufallszahlen so transformiert, dass das zufällige Experiment auch richtig modelliert wird. Wenn wir zum Beispiel ein Würfelspiel simulieren wollen, dann ist es wichtig, dass wir die Zufallszahlen so transformieren, dass die Zahlen 1, 2, 3, 4, 5 und 6 mit möglichst gleicher Wahrscheinlichkeit „gezogen“ werden, wie wir das im Grundkurs (Kurs 1) ja auch bereits gemacht haben.

Nehmen wir nun an, wir haben n Simulationsschritte durchgeführt. Bei jedem Simulationsschritt können wir den Computer bestimmen lassen, ob A eingetroffen ist oder nicht, d.h. wir stellen eine mathematische Bedingung dafür auf, dass A eingetreten ist und diese Bedingung wird nach jedem Simulationsschritt vom Computer überprüft. Dann können wir die **relative**

Häufigkeit dafür bestimmen, dass A eingetreten ist, indem wir zählen, bei wie vielen Simulationsschritten das Ereignis A eingetroffen ist und diese Anzahl dann durch die Gesamtanzahl der Simulationsschritte teilen:

$$h_n(A) = \frac{\text{Anzahl der Simulationsschritte, in denen } A \text{ eingetreten ist}}{n}.$$

Dadurch bekommen wir für alle $n \in \mathbb{N}$ einen Wert

$$0 \leq h_n(A) \leq 1,$$

der für große n näherungsweise gleich der theoretischen Wahrscheinlichkeit dafür ist, dass A eintritt, also gleich $P(A)$. Für große $n \in \mathbb{N}$ gilt also:

$$h_n(A) \approx P(A).$$

Doch was heißt das genau? Angenommen, wir wiederholen einen Versuch 10000mal. Erhalten wir dann mit $h_{10000}(A)$ zwangsläufig einen sehr guten Näherungswert für $p(A)$? Welche Beziehung herrscht also zwischen Theorie (Wahrscheinlichkeitstheorie) und Praxis (angewandter Statistik/Simulation)? Sagt die Theorie die Praxis richtig voraus und kann umgekehrt (und darauf kommt es uns hier ja an) aus den Beobachtungen (hier: Simulationen) auf die theoretischen Werte geschlossen werden?

Bei tatsächlichen zufälligen Versuchen gilt das **Gesetz der Großen Zahlen**, und ein solches gibt es auch für simulierte Experimente: Die relativen Häufigkeiten nähern sich mit immer größerer Wahrscheinlichkeit immer näher an die wahren (theoretischen) Wahrscheinlichkeiten an. Dieser Sachverhalt erlaubt es uns, von relativen Häufigkeiten (die wir aus echten zufälligen Experimenten oder eben stochastischen Simulationen erlangen) auf theoretische Wahrscheinlichkeiten zu schließen - allerdings immer verbunden mit einem Rest Unsicherheit. Diese Unsicherheit wird umso kleiner, je mehr Experimente/Simulationen wir durchführen.

Diese Methode zur näherungsweisen Bestimmung von Wahrscheinlichkeiten bezeichnet man als **Monte-Carlo-Simulation**. Der Begriff „Monte-Carlo-Simulation“ entstand, wie bereits im Grundkurs (Kurs 1) erwähnt, in den 40er Jahren, als man im Zusammenhang mit dem Bau der Neutronenbombe die Simulation von Zufallsprozessen erstmals in größerem Stil einsetzte, um die Wechselwirkung von Neutronen mit Materie theoretisch vorherzusagen.

VON NEUMANN benutzte dabei die Technik der Simulation, um mathematische Funktionen, die nicht explizit programmiert werden konnten, abzubilden. Das Codewort für diese geheimen Arbeiten war „Monte Carlo“. Die Bezeichnung ist eine Anspielung auf den für Glücksspiele bekannten Stadtteil von Monaco, da die Grundlage des Verfahrens Zufallszahlen sind, wie man sie auch mit einem Roulette-Rad erzeugen könnte. Schon damals wurde eine ganze Reihe von grundlegenden Verfahren entwickelt, und heute zählen Monte-Carlo-Simulationen zu den Standardverfahren und lassen sich auf viele naturwissenschaftliche, finanzmathematische, technische und medizinische Probleme mit großem Erfolg anwenden.

Seit dieser Zeit ist also die dort verwendete Simulationsmethode als Monte-Carlo-Simulation bekannt. Dies bedeutet aber nicht, dass diese Techniken nicht schon vorher verwendet wurden. Der Ursprung der Methode kann bis in die Mitte der 19. Jahrhunderts zurückverfolgt werden. Damals wurden solche Simulationsmethoden beispielsweise zur näherungsweisen Berechnung der Zahl π eingesetzt. (Diese Anwendung werden wir in diesem Kurs genauer kennenlernen.) Zu Beginn des 20. Jahrhunderts lieferten Simulationsverfahren Fortschritte bei der Untersuchung der Boltzmann-Gleichung.

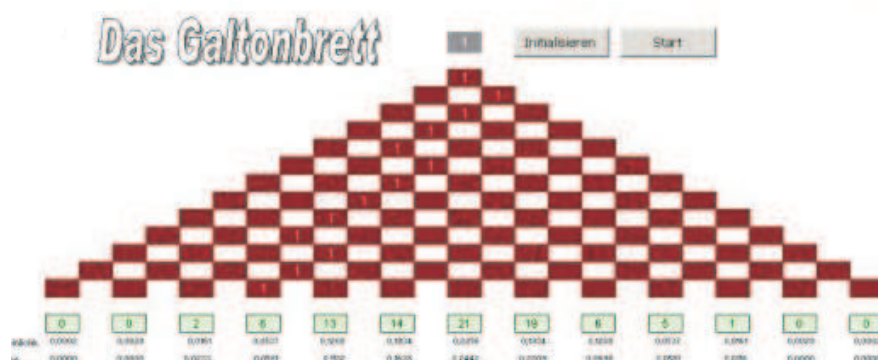
Zur Hintergrundinformation: Die Boltzmann-Gleichung erklärt, wie das Verhalten von Gasen von der ständigen Bewegung der Atome und Moleküle abhängt. Ihre große Bedeutung liegt in der Beschreibung zahlreicher naturwissenschaftlicher Phänomene und technischer Gerätschaften, bei denen Gase eine wichtige Rolle spielen: in allen Dampf- oder Verbrennungsmaschinen, bei den zahllosen Reaktionen zwischen Gasen, die zur chemotechnischen Herstellung moderner Medikamente, Kunststoffe usw. verwendet werden, beim Verständnis des Wetters und auch für die Erklärung der gewaltigen Vorgänge in der Sonne, den Sternen und fernen Galaxien.

Kapitel 2

Simulation der Binomialverteilung

Simulation 1: Das Galton-Brett

In dem bereitgestellten Excel-Programm sind 13 Reihen in einer Dreiecksform angeordnet. In der obersten Reihe befindet sich ein Kästchen, in der zweitobersten Reihe befinden sich zwei Kästchen (gegenüber dem ersten Kästchen versetzt), in der drittobersten drei Kästchen usw. Es kommt also immer ein Kästchen dazu, so dass wir in der untersten Reihe 13 Kästchen haben (schau dir das in dem Bild an, es geht um die braunen Kästchen):



Das Galton-Brett in Excel

Mit diesem Programm wird das berühmte Galton'sche¹ Nagelbrett simuliert, bei dem Kugeln zufällig durch in ein Brett geschlagene Nägel fallen.

Bei unserem Simulationsexperiment „wirft“ man die Kugel in das Feld der obersten Reihe und schaut, was passiert: Von einer Reihe in die nächstuntere hat die Kugel jeweils zwei Möglichkeiten: entweder sie „fällt“ nach links oder aber nach rechts. Insgesamt „muss sich die Kugel also 12mal entscheiden“, zu welcher der beiden Seite sie „fällt“.

Der Verlauf der Kugel wird im Programm durch die ausgefüllten Kreise angezeigt. Sprich: Überall dort, wo ein ausgefüllter Kreis zu finden ist, ist die Kugel „durchgelaufen“.

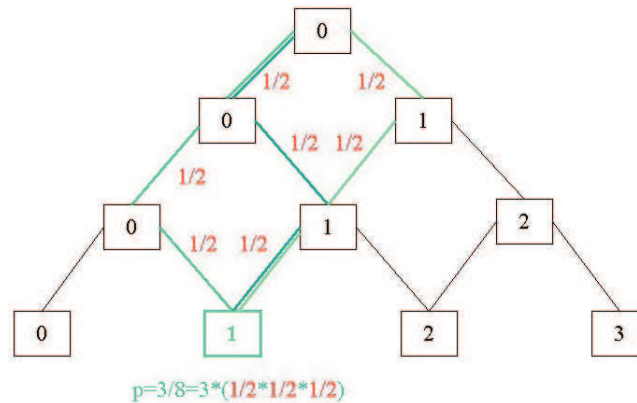
Gruppenaufgabe:

Bevor du die Simulation durchführst: Glaubst du, dass in der ganz untersten Reihe nach einiger Zeit in allen Feldern ungefähr gleich viele Kugeln ankommen oder wird es Unterschiede geben? Wo werden tendenziell die meisten Kugeln ankommen?

Male dir mal ein Galton-Brett mit nur vier Reihen auf und fasse das als Baumdiagramm mit der Elementarwahrscheinlichkeit $p = \frac{1}{2}$ auf. Berechne jetzt die Wahrscheinlichkeit, mit der die Kugeln in den vier Kästchen der vierten Reihe ankommen.

Ich habe euch das zur Kontrolle mal aufgezeichnet:

¹Sir Francis Galton, geboren am 16.02.1822 in Birmingham, gestorben am 17.01.1911 in Haslemere, war britischer Vererbungsforscher, Mitbegründer der Eugenik (von griechisch *eugenes*: wohlgeboren; die historische Bezeichnung für die Anwendung der Erkenntnisse der Humangenetik auf Bevölkerungsentwicklungen; der Begriff ist sehr negativ besetzt und beschreibt eine primitive biologische Lehre, der ausschließlich das Prinzip der natürlichen Auslese zu Grunde liegt; er wurde zudem von den Nazis für ihre „Rassenlehre“ missbraucht) und der Zwillingsforschung und beschäftigte sich mit der Vererbung geistiger Eigenschaften.



schemantisches Galton-Brett mit vier Reihen

Die Zahlen in den Kästchen bedeuten, wie oft die Kugel vorher nach rechts gefallen ist um zu diesem Kästchen zu gelangen. Wie groß ist die Wahrscheinlichkeit zu dem Kästchen in der vierten Reihe mit der „1“ zu gelangen?

Nun, es gibt drei Wege dorthin zu gelangen. Jeder einzelne Weg hat die Wahrscheinlichkeit:

$$p = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^3 = \frac{1}{8},$$

denn in jedem der drei Schritte beträgt die Wahrscheinlichkeit $\frac{1}{2}$. Da wir drei Wege haben und die Wahrscheinlichkeit jedes Weges gleich $\frac{1}{8}$ ist, beträgt die Wahrscheinlichkeit einen dieser drei Wege zu benutzen und auf diesem zum Feld „1“ in der vierten Reihe zu gelangen, gerade:

$$P(X_4 = 1) = 3 \cdot \frac{1}{8} = \frac{3}{8} = 0,375 = 37,5\%,$$

wobei X_4 die Anzahl der Bewegungen der Kugel nach rechts bis zur vierten Reihe beschreibt.

Jetzt wollen wir uns mal allgemein für dreizehn Reihen die theoretischen Wahrscheinlichkeiten dafür überlegen, mit der die Kugeln in den einzelnen Kästchen der letzten, also 13ten, Reihe ankommen. Allerdings sind die Wahrscheinlichkeiten nicht ganz einfach zu berechnen und führen zu Überlegungen, die eher Thema der späten Oberstufe sind (wenn sie denn überhaupt behandelt werden). Wir wollen an dieser Stelle aber dennoch das Prinzip verstehen, und ich gebe euch dann anschließend die genauen Ergebnisse einfach nur an.

Danach allerdings wollen wir uns dann noch einen alternativen Zugang zu diesem Problem anschauen, der thematisch zu der Unter- und Mittelstufe zurückführt (das Pascal'sche Dreieck).

Überlegen wir uns mal die Wahrscheinlichkeit dafür, dass die Kugel ganz links ankommt. Dann muss sie sich bei den beiden Möglichkeiten „links“ (im Folgenden abgekürzt mit l) und „rechts“ (im Folgenden abgekürzt mit r) immer für „links“ entschieden haben, d.h. wir erhalten die Folge:

lllllllllll.

Die Wahrscheinlichkeit, dass l kommt, ist jedes Mal gleich $p = \frac{1}{2}$. Die Wahrscheinlichkeit, dass 12 mal hintereinander l kommt, ist nach der Pfadregel:

$$P(X_{13} = 0) = P(\text{„12 mal l“}) = P(\text{„lllllllllll“}) = \left(\frac{1}{2}\right)^{12}.$$

Ebenso gilt natürlich für die Wahrscheinlichkeit, dass die Kugel im Feld ganz rechts ankommt:

$$P(X_{13} = 12) = P(\text{„12 mal r“}) = P(\text{„rrrrrrrrrrr“}) = \left(\frac{1}{2}\right)^{12}.$$

Nun überlegen wir uns die Wahrscheinlichkeit, dass die Kugel im zweiten Feld von links ankommt. Dazu muss die Kugel 11mal nach links und einmal nach rechts gerollt sein. Eine Möglichkeit dafür wäre also die Folge „rlllllllllll“. Eine andere Möglichkeit wäre zum Beispiel „lllllrlllll“. Wie viele solcher Möglichkeiten gibt es? So viele Möglichkeiten, wie es Positionen gibt, an denen das eine „r“ (neben den 11 „l“s) stehen kann, also 12 Stück. Für jede dieser Möglichkeiten beträgt die Wahrscheinlichkeit wieder $\left(\frac{1}{2}\right)^{12}$. Daher gilt:

$$P(X_{13} = 1) = P(\text{„11 mal l, 1 mal r“}) = 12 \cdot \left(\frac{1}{2}\right)^{12}.$$

Ebenso gilt natürlich für die Wahrscheinlichkeit, dass die Kugel im zweiten Feld von rechts ankommt:

$$P(X_{13} = 11) = P(\text{„11 mal r, 1 mal l“}) = 12 \cdot \left(\frac{1}{2}\right)^{12}.$$

Nun überlegen wir uns die Wahrscheinlichkeit, dass die Kugel im dritten Feld

von links ankommt. Dazu muss die Kugel 10mal nach links und 2mal nach rechts gerollt sein. Eine Möglichkeit dafür wäre also die Folge „rrlllllllll“. Eine andere Möglichkeit wäre „llllrlllrl“. Wie viele solcher Möglichkeiten gibt es? So viele Möglichkeiten, wie es Positionen gibt, an denen die beiden r 's stehen können. Für das erste r gibt es zwölf Möglichkeiten. Dann verbleiben für das zweite r noch elf Möglichkeiten, denn es kann ja nicht an der gleichen Stelle wie das erste r stehen. Daher gibt es zunächst einmal $12 \cdot 11$ Möglichkeiten, wo die beiden r 's stehen können. Allerdings haben wir dabei jede der Möglichkeiten doppelt gezählt. Warum? Nun ja, wir können die beiden r 's ja nicht unterscheiden! Bisher haben wir jedoch zum Beispiel den Pfad „llllrlllrl“ zweimal gezählt, da wir so gerechnet haben, als könnten wir die beiden r 's unterscheiden. Um diesen Fehler zu korrigieren, müssen wir die gefundene Anzahl noch durch 2 teilen. Insgesamt gibt es also $\frac{12 \cdot 11}{2}$ Möglichkeiten, zehn mal ein l und zweimal ein r zu wählen. Für jede dieser Möglichkeiten beträgt die Wahrscheinlichkeit wieder $\left(\frac{1}{2}\right)^{12}$. Daher gilt für die Wahrscheinlichkeit, dass die Kugel im dritten Feld von links ankommt:

$$P(X_{13} = 2) = P(\text{„10 mal l, 2 mal r“}) = \frac{12 \cdot 11}{2} \cdot \left(\frac{1}{2}\right)^{12}.$$

Ebenso gilt natürlich für die Wahrscheinlichkeit, dass die Kugel im dritten Feld von rechts ankommt:

$$P(X_{13} = 10) = P(\text{„10 mal r, 2 mal l“}) = \frac{12 \cdot 11}{2} \cdot \left(\frac{1}{2}\right)^{12}.$$

Allgemein: Wie hoch ist die Wahrscheinlichkeit, dass die Kugel im $(k+1)$ -ten Feld von links ankommt? Dazu muss die Kugel $(12-k)$ -mal nach links und k -mal nach rechts gerollt sein. Eine Möglichkeit dafür wäre also (mit $k=5$) die Folge „rrrrrlllllll“. Eine andere Möglichkeit wäre „lrrllrlllrl“. Wie viele solcher Möglichkeiten gibt es? So viele Möglichkeiten, wie es Positionen gibt, an denen die k r 's stehen können. Für das erste r gibt es zwölf Möglichkeiten. Dann verbleiben für das zweite r noch elf Möglichkeiten, denn es kann ja nicht an der gleichen Stelle wie das erste r stehen. Für das dritte r gibt es zehn Möglichkeiten, denn es muss ja an einer anderen Stelle als die beiden anderen r 's stehen. Für das k -te r gibt es somit $12-k+1$ Möglichkeiten. Daher gibt es zunächst einmal $12 \cdot 11 \cdot \dots \cdot (12-k+1)$ Möglichkeiten, wo die k r 's stehen können. Allerdings haben wir dabei jede der Möglichkeiten mehrfach gezählt. Warum? Nun ja, wir können die k r 's ja nicht unterscheiden! Um diesen Fehler zu korrigieren, müssen wir die gefundene Anzahl noch

durch die Anzahl der möglichen Vertauschungen der r 's untereinander teilen. Wie viele solcher Vertauschungen gibt es? Das erste r kann an k Positionen stehen, das zweite r an $k - 1$ Positionen, usw. Somit gibt es $k \cdot (k - 1) \cdot \dots \cdot 1$ Möglichkeiten die r 's untereinander zu vertauschen.

Durch diese Anzahl müssen wir unsere $12 \cdot 11 \cdot \dots \cdot (12 - k + 1)$ Möglichkeiten teilen. Insgesamt gibt es also $\frac{12 \cdot 11 \cdot \dots \cdot (12 - k + 1)}{k \cdot (k - 1) \cdot \dots \cdot 1}$ Möglichkeiten, $(12 - k)$ -mal ein l und k -mal ein r zu wählen. Die Mathematiker haben dafür ein eigenes Symbol definiert, den sogenannten **Binomialkoeffizienten**. Es gilt:

$$\binom{12}{k} = \frac{12 \cdot 11 \cdot \dots \cdot (12 - k + 1)}{k \cdot (k - 1) \cdot \dots \cdot 1}.$$

Für jede dieser Möglichkeiten beträgt die Wahrscheinlichkeit wieder $\left(\frac{1}{2}\right)^{12}$. Daher gilt für die Wahrscheinlichkeit, dass die Kugel im $(k + 1)$ -ten Feld von links ankommt:

$$P(X_{13} = k) = P(„(12 - k) \text{ mal } l, k \text{ mal } r“) = \binom{12}{k} \cdot \left(\frac{1}{2}\right)^{12}.$$

Ebenso gilt natürlich für die Wahrscheinlichkeit, dass die Kugel im $(k + 1)$ -ten Feld von rechts ankommt:

$$P(X_{13} = 12 - k) = P(„k \text{ mal } l, (12 - k) \text{ mal } r“) = \binom{12}{k} \cdot \left(\frac{1}{2}\right)^{12}.$$

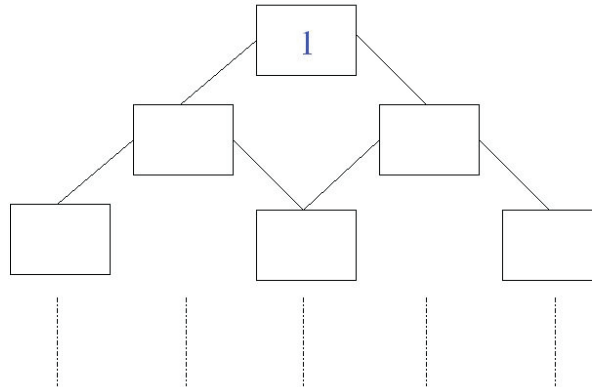
Das Grundprinzip ist, hoffe ich, klar geworden. Grob gesagt lautet es so: Die Wahrscheinlichkeit für jeden einzelnen Pfad ist gleich groß, nämlich gleich $\left(\frac{1}{2}\right)^{12}$, da es zwölf voneinander unabhängige Entscheidungsmöglichkeiten zwischen zwei Ereignissen („links“ oder „rechts“) gibt, die beide mit der Wahrscheinlichkeit $\frac{1}{2}$ eintreten. Man schaut sich für jede Anzahl k von r 's an, auf wie viele Arten man diese r 's anordnen kann, ohne dass man manche Anordnungen doppelt zählt. Diese Anzahl entspricht dann der Anzahl der Pfade, bei denen sich eine Kugel k mal für „rechts“ und $(12 - k)$ -mal für „links“ entscheidet. Wenn sie sich k mal für rechts entscheidet, landet die Kugel in der letzten Reihe im $(k + 1)$ -ten Kästchen von links. (Ist dir das klar?)

Das führt dann zu einer allgemeinen Formel. Stochastische Versuchsreihen dieser Art bezeichnet man als **Bernoulli-Kette**, die entstehende Wahr-

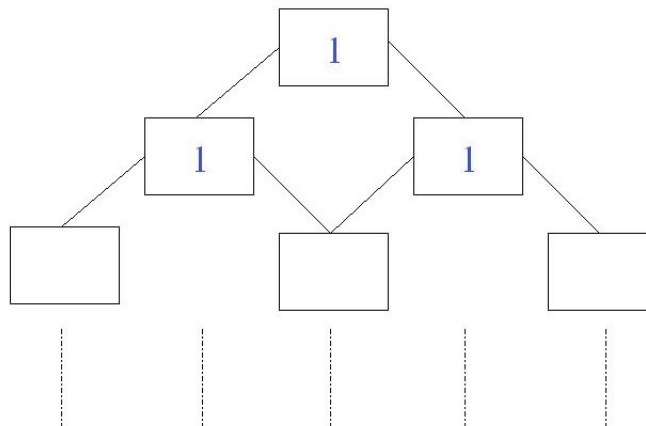
scheinlichkeitsverteilung heißt **Binomialverteilung**. Da dies aber eher Stoff der Oberstufe ist, beschäftigen wir uns an dieser Stelle nicht detaillierter mit dieser Wahrscheinlichkeitsverteilung und ich gebe stattdessen die theoretischen Wahrscheinlichkeiten, mit der die Kugel in den einzelnen Kästchen der letzten Reihe landet, zunächst einmal nur an:

$$\begin{aligned}
P(\text{„die Kugel landet im ganz linken Kästchen“}) &= 0,000244141, \\
P(\text{„die Kugel landet im zweiten Kästchen von links“}) &= 0,002929688, \\
P(\text{„die Kugel landet im dritten Kästchen von links“}) &= 0,016113281, \\
P(\text{„die Kugel landet im vierten Kästchen von links“}) &= 0,053710938, \\
P(\text{„die Kugel landet im fünften Kästchen von links“}) &= 0,120849609, \\
P(\text{„die Kugel landet im sechsten Kästchen von links“}) &= 0,193359375, \\
P(\text{„die Kugel landet im mittleren Kästchen“}) &= 0,225585938, \\
P(\text{„die Kugel landet im sechsten Kästchen von rechts“}) &= 0,193359375, \\
P(\text{„die Kugel landet im fünften Kästchen von rechts“}) &= 0,120849609, \\
P(\text{„die Kugel landet im vierten Kästchen von rechts“}) &= 0,053710938, \\
P(\text{„die Kugel landet im dritten Kästchen von rechts“}) &= 0,016113281, \\
P(\text{„die Kugel landet im zweiten Kästchen von rechts“}) &= 0,002929688, \\
P(\text{„die Kugel landet im ganz rechten Kästchen“}) &= 0,000244141.
\end{aligned}$$

So, jetzt aber werdem wir uns einen ganz anderen Ansatz überlegen, und dieser wird uns die Sachlage einmal von einem neuen Standpunkt aus durchleuchten. Schauen wir uns das Galton-Brett einmal schematisch an: Wie viele Möglichkeiten gibt es in das allererste Feld zu gelangen? Klar, natürlich nur eine Möglichkeit! Schreiben wir dies doch einmal in das erste Kästchen rein:

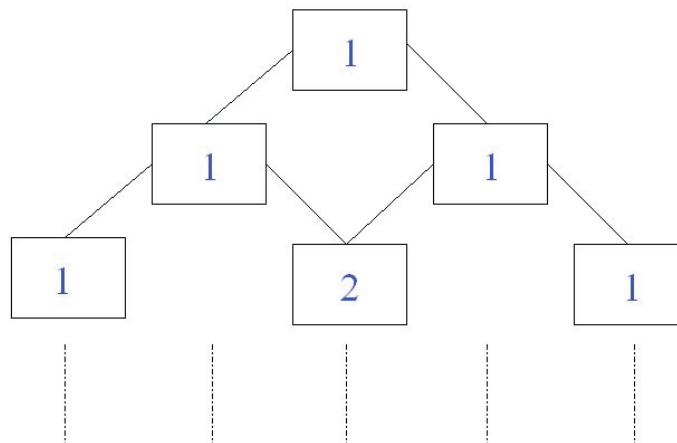


Wie viele Möglichkeiten gibt es in eines der beiden Felder in der zweiten Reihe zu gelangen? Ebenfalls jeweils eine Möglichkeit! Wir haben also die Situation:

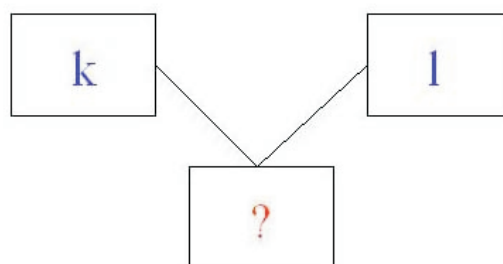


Wie viele Möglichkeiten erhält man nun für die dritte Reihe? Nun, für die beiden äußeren Felder gibt es wieder nur eine Möglichkeit. Aber wie sieht es mit dem mittleren Feld aus? Man kann ja vorher auf dem Feld „links oben“ oder auf dem Feld „rechts oben“, jeweils von dem mittleren Feld der dritten Reihe aus betrachtet, gewesen sein. In beiden Fällen gab es vorher eine Möglichkeit dorthin gelangt zu sein. Und von beiden Feldern aus kann man auf das mittlere der drei Felder der dritten Reihe kommen. Daher gibt es

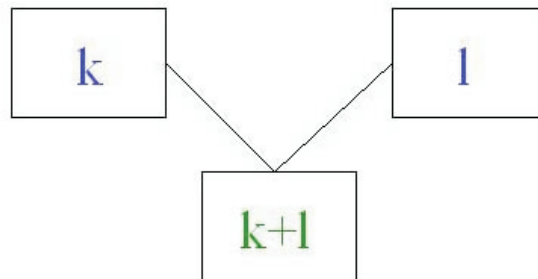
insgesamt $1 + 1 = 2$ Möglichkeiten auf das mittlere Feld der dritten Reihe zu gelangen. Wir erhalten also:



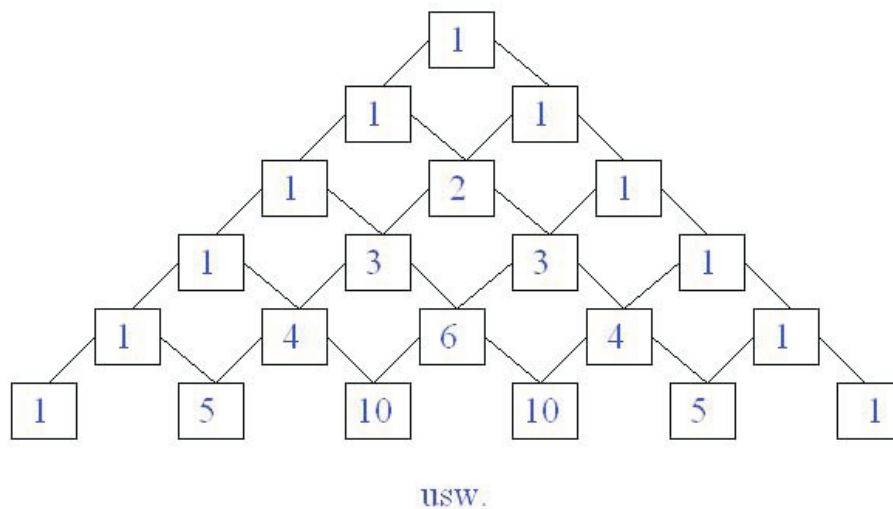
Und jetzt schauen wir uns mitten im Galton-Brett irgendein Feld in irgendeiner Reihe an. Auf wie viele Arten kann man dorthin gelangen? Nun, vorher muss auf einem der beiden Felder gewesen sein, die sich direkt „links über“ oder „rechts über“ diesem Feld befinden. Man hat also die folgende Situation:



Die Anzahl der Möglichkeiten auf unser beliebig gewähltes Feld zu gelangen, ist also gleich der Anzahl der Möglichkeiten auf das Feld „unmittelbar links darüber“ zu gelangen (und dann nach rechts zu gehen) plus der Anzahl der Möglichkeiten auf das Feld „unmittelbar rechts darüber“ zu gelangen (und dann nach links zu gehen). Somit haben wir:



Die Zahl in einem beliebigen Kästchen ist also gleich der Summe der beiden Zahlen in den Kästchen, die sich unmittelbar links und rechts über diesem Kästchen befinden. erinnert euch das an etwas? Genau! Dies ist genau die (rekursive) Vorschrift für das Pascalsche Dreieck! Wir sehen also, dass für die Anzahl der Möglichkeiten, zu einem Kästchen zu gelangen, das folgende Schema gültig ist:



Andererseits hatten wir hergeleitet, dass diese Zahlen in der $(n+1)$ -ten Reihe gerade die Binomialkoeffizienten

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} \quad (k = 0, 1, \dots, n)$$

sind. Auf diese Weise können wir unmittelbar die Werte für die Binomialkoeffizienten durch einen Vergleich ablesen, etwa:

$$\binom{5}{2} = 10.$$

(Und in der Tat gilt ja: $\binom{5}{2} = \frac{5 \cdot 4}{1 \cdot 2} = 10$.)

Weiterhin wisst ihr vielleicht aus der Schule, dass die Zahlen des Pascalschen Dreiecks genau die Vorfaktoren vor den n -ten Potenzen der Summe zweier Zahlen sind. So gilt etwa:

$$\begin{aligned}(a+b)^4 &= 1 \cdot a^4 \cdot b^0 + 4 \cdot a^3 \cdot b^1 + 6 \cdot a^2 \cdot b^2 + 4 \cdot a^1 \cdot b^3 + 1 \cdot a^0 \cdot b^4 \\ &= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4.\end{aligned}$$

Die Vorfaktoren sind die Zahlen des Pascalschen Dreiecks in der $(4+1)$ -ten, also 5-ten Zeile. Gleichzeitig stehen dort aber, wie wir jetzt wissen, die Binomialkoeffizienten:

$$\binom{4}{0} \quad , \quad \binom{4}{1} \quad , \quad \binom{4}{2} \quad , \quad \binom{4}{3} \quad , \quad \binom{4}{4}.$$

Daher gilt:

$$(a+b)^4 = \binom{4}{0} \cdot a^4 \cdot b^0 + \binom{4}{1} \cdot a^3 \cdot b^1 + \binom{4}{2} \cdot a^2 \cdot b^2 + \binom{4}{3} \cdot a^1 \cdot b^3 + \binom{4}{4} \cdot a^0 \cdot b^4$$

oder allgemeiner:

$$(a+b)^n = \binom{n}{0} \cdot a^n \cdot b^0 + \binom{n}{1} \cdot a^{n-1} \cdot b^1 + \dots + \binom{n}{n-1} \cdot a^1 \cdot b^{n-1} + \binom{n}{n} \cdot a^0 \cdot b^n,$$

was die Aussage des sehr wichtigen „Binomischen Lehrsatzes“ ist.

Gruppenaufgabe:

Führe die Simulation jetzt ein paar Mal durch. Vergleiche dabei die relativen Häufigkeiten, die sich aus dem Simulationsexperiment ergeben, mit den gerade angegebenen theoretischen Wahrscheinlichkeiten. Dies kannst du auch in dem Excel-Blatt mit Hilfe des Diagramms optisch verfolgen. Ändere jetzt mal den Wert p , also den Wert für die Wahrscheinlichkeit, mit der die Kugel nach links rollt. Wie ändern sich jetzt die theoretischen Wahrscheinlichkeiten? Wird das auch in der Simulation deutlich? Versuche auch den hinterliegenden Visual-Basic-Code zu verstehen.

Gruppenaufgabe:

Wir versuchen nun zusammen, ein solches Galton-Brett mit elementaren Tabellenfunktionen selber kurz in Excel zu erstellen.

Wann tauchen jetzt solche Binomialverteilungen auf? Dazu schauen wir uns mal ein Beispiel an:

Beispiel:

Ein Multiple-Choice-Test besteht aus zwölf Fragen mit jeweils vier Antwortmöglichkeiten, von denen jeweils genau eine richtig ist. Kreuzt man in diesem Test nach Zufall an, etwa weil man nicht vorbereitet und daher aufs bloße Raten angewiesen ist, so „trifft“ man bei jeder Frage genau mit einer Wahrscheinlichkeit $p = \frac{1}{4} = 0,25$ die richtige Antwort. Es interessiert uns nun natürlich die Anzahl X der richtig beantworteten Fragen. Dies entspricht genau der obigen Situation. Hier kann jede Frage falsch oder richtig sein, unabhängig von allen anderen Fragen, im Galton-Brett konnte die Kugel an jedem Hindernis nach links oder rechts rollen, unabhängig von allen anderen Hindernissen. Wieder haben wir also eine sogenannte Bernoulli-Kette, d.h. X genügt einer Binomialverteilung. Diesmal haben wir aber nicht die (Erfolgs-)Wahrscheinlichkeit $p = 0,5$ (die Wahrscheinlichkeit nach links oder rechts zu rollen, war ja jeweils 0,5), sondern die (Erfolgs-)Wahrscheinlichkeit

$p = 0,25$ (die Antwort richtig zu haben, d.h. für die Kugel: nach links zu rollen) bzw. die (Misserfolgs-)Wahrscheinlichkeit $1 - p = 0,75$ (die Antwort falsch zu haben, d.h. für die Kugel: nach rechts zu rollen).

Gruppenaufgabe:

Nehmen wir an, man besteht den Test, wenn man mindestens die Hälfte der Fragen richtig beantwortet hat. Bestimme mit der stochastischen Simulation des Galton-Bretts approximativ (also näherungsweise) die Wahrscheinlichkeit für dieses Ereignis. Lohnt es sich also zu raten? ;-)

Kapitel 3

Klassische Monte-Carlo-Simulation

Simulation 2: Monte-Carlo-Simulation zur näherungsweisen Bestimmung der Kreiszahl π

Wir möchten nun mit einer (ja, der wohl bekanntesten) Monte-Carlo-Simulation einen Näherungswert für die Kreiszahl π erhalten. Dazu müssen wir uns zunächst einmal vergegenwärtigen, dass man Flächen auch als Wahrscheinlichkeiten deuten kann. Man nennt diese Wahrscheinlichkeiten häufig (jedenfalls in der Schule) „geometrische Wahrscheinlichkeiten“. Dieses Konzept möchte ich kurz vorstellen:

Zunächst im Eindimensionalen: Es sei I ein endliches Intervall, zum Beispiel $I = [0, 2]$. Wir bezeichnen die Länge eines solchen Intervalls mit $|I|$. (Damit gilt zum Beispiel: $|[0, 2]| = 2$.) Ist nun A ein Teilintervall von I mit der Länge $|A|$, so betrachten wir die durch

$$P(A) := \frac{|A|}{|I|}$$

definierte Zahl.

Hier besteht die zugrundeliegende Menge, also das Intervall I , aus unendlich vielen Punkten. Auf solchen Mengen ist es sehr viel schwieriger zu

definieren, was erlaubte Ereignisse und Wahrscheinlichkeiten sein sollen. Insbesondere können wir keine Funktion $P : \mathcal{P}(I) \rightarrow [0, 1]$ finden (die also auf **allen** Teilmengen von I definiert ist), welche all den Eigenschaften genügt, die wir von einem Wahrscheinlichkeitsmaß erwarten. Wir können aber eine Abbildung P finden, die auf der Menge aller möglichen endlichen Vereinigung von Teilintervallen von I durch

$$P(A_1 \cup A_2 \cup \dots \cup A_n) := \frac{|A_1|}{|I|} + \frac{|A_2|}{|I|} + \dots + \frac{|A_n|}{|I|}$$

definiert ist, wenn die Intervalle paarweise leeren Durchschnitt besitzen (wenn also $A_i \cap A_j = \emptyset$ gilt für $i \neq j$). Dies kann man sogar für unendliche Vereinigungen von Intervallen erweitern, aber das lassen wir an dieser Stelle. Es soll nur ein kurzer Einblick in die komplizierte Welt der Wahrscheinlichkeitstheorie auf unendlichen Mengen sein!

Dieses Konzept kann man natürlich ins Zweidimensionale (und dann auch ins Höherdimensionale) übertragen. Machen wir das mal im Zweidimensionalen, denn da brauchen wir es gleich: Es sei G ein Teilgebiet der Zahlenebene (zum Beispiel ein Quadrat) mit dem Flächeninhalt $|G|$. Dann wird durch

$$P(A) = \frac{|A|}{|G|}$$

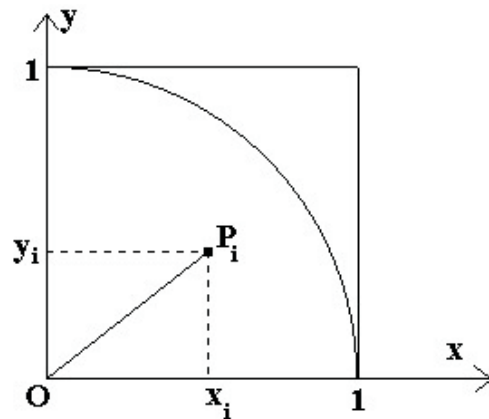
auf **gewissen** (nicht allen!) Teilmengen A von G eine Funktion definiert, die man ebenfalls wie oben auf paarweise disjunkte Vereinigungen von gewissen Teilmengen erweitern kann. Hierbei ist $|A|$ der Flächeninhalt von A . Erfasst werden dadurch aber immerhin „alle Teilmengen von G , die wir uns vorstellen können“, denen wir also auf vernünftige Weise einen Flächeninhalt zuordnen können, **sehr salopp** gesagt. Dabei besitzt jeder Kurvenzug (zum Beispiel ein Geradenstück) die Wahrscheinlichkeit Null, da solche Kurvenzüge den Flächeninhalt Null besitzen. Daraus folgt jedoch nicht, dass kein Punkt auf diesem Kurvenzug als Ereignis eintreten kann! Die Menge der Punkte auf einem solchen Kurvenzug ist aber im Vergleich zur Gesamtmenge G vernachlässigbar klein. Daher wird das Versuchsergebnis höchst selten (d.h. fast nie, also mit keiner positiven Wahrscheinlichkeit) auf diesem Kurvenzug liegen. Klingt unglaublich und seltsam? Dem stimme ich zu! Aber mit etwas mehr mathematischer Erfahrung kann man dies durchaus „verstehen“.

Nun bleiben wir im Zweidimensionalen und betrachten das Einheitsquadrat

$$G = [0, 1]^2 = [0, 1] \times [0, 1] = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1\}.$$

Offenbar gilt für die Fläche dieses Quadrates: $|G| = 1$.

Weiterhin sei A der Viertelkreis um 0 mit Radius $r = 1$:



Viertelkreis A im Einheitsquadrat G

Für die Fläche eines Kreises mit Radius r gilt bekanntlich:

$$\text{Fläche}_{\text{Kreis}} = \pi \cdot r^2.$$

Daher gilt für die Fläche des Viertelkreises A mit Radius 1:

$$|A| = \frac{1}{4} \pi r^2 = \frac{\pi}{4}$$

und damit auch:

$$P(A) = \frac{|A|}{|G|} = \frac{\pi}{4}.$$

Wir erzeugen nun n Zufallswerte aus $[0, 1] \times [0, 1]$, also Paare von Zufallszahlen, die jeweils zwischen 0 und 1 liegen und lassen den Computer bestimmen,

wie viele dieser Zufallswerte im Viertelkreis A liegen.

Nehmen wir jetzt mal an, dass k der n zufälligen Paare in A liegen, d.h. dass für die relative Häufigkeit folgendes gilt:

$$h_n(A) = \frac{k}{n}.$$

Dann gilt für große n approximativ (ungefähr), da sich für große n die relative Häufigkeit der theoretischen Wahrscheinlichkeit annähert:

$$\frac{\pi}{4} = P(A) \approx h_n(A) = \frac{k}{n},$$

also, wenn man es nach π umstellt:

$$\pi \approx 4 \cdot \frac{k}{n}.$$

Damit haben wir in $4 \cdot \frac{k}{n}$ einen Näherungswert für π gefunden, der tendenziell umso genauer wird, je größer n wird.

Aufgabe 1:

Versuche nun in einem Excel-Blatt (und mit Hilfe der Zufallszahlen von Excel) einen Näherungswert für π zu bestimmen, indem du ein kleines stochastisches Simulationsprogramm schreibst und dann $n = 10000$ Simulationsschritte durchführst.

(TIPP: Ein Punkt $(x_i, y_i) \in [0, 1] \times [0, 1]$ liegt genau dann innerhalb des Viertelkreises A (oder auf dem Kreisrand), wenn $x_i^2 + y_i^2 \leq 1$ gilt. Mach dir das bitte mit Hilfe des Satzes von Pythagoras klar. Wenn du nicht selber vorankommst, dann kannst du dir auch das bereits angefertigte Excel-Programm zum Vorbild nehmen oder es einfach benutzen.)

Natürlich lassen sich auch kompliziertere Flächeninhalte und Volumina näherungsweise mit solchen Monte-Carlo-Methoden berechnen.

Aufgabe 2:

Berechne mit einer Monte-Carlo-Simulation näherungsweise den Flächeninhalt zwischen dem Graphen der Funktion $f(x) = -\frac{1}{4}x^2 + 1$ und der x -Achse. Betrachte dazu die Obermenge $G = [-2, 2] \times [0, 1]$ und überlege dir, wann ein Punkt aus G zwischen dem Graphen und der x -Achse liegt. Eine Skizze ist hier sehr hilfreich! Beachte, dass du Zufallszahlen aus $[0, 1]$ zunächst auf $[-2, 2]$ transformieren musst. Ist nun X eine Zufallszahl aus $[-2, 2]$ und Y eine Zufallszahl aus $[0, 1]$, dann musst du $f(X)$ und Y vergleichen.

Zusatzaufgabe:

Bilde jetzt mal den Mittelwert der $f(X_i)$, wobei die X_i die Zufallszahlen aus $[-2, 2]$ und multipliziere den Wert wiederum mit 4. Du hast gerade zwei mögliche Monte-Carlo-Methoden zur Berechnung des Flächeninhalts unter einem Funktionsgraphen kennengelernt!

Wir wollen uns diesen letzten Fall der Monte-Carlo-Simulation zur näherungsweisen Berechnung eines bestimmten Integrals (auf einem beschränkten Intervall) noch einmal näher anschauen. Hierbei verwenden wir die einfachste Methode, indem wir gleichverteilte Zufallszahlen erzeugen (also Zufallszahlen, die auf einem gegebenen Intervall, anschaulich gesagt, gleichmäßig verteilt sind). Im Allgemeinen ist es besser andere Verteilungen zuzulassen und diese dann anschließend wieder „rauszurechnen“, aber das würde hier zu weit führen, und in diesem einfachsten Fall wird die Idee der Monte-Carlo-Simulation bereits sehr gut deutlich.

Berechnen wollen wir also für eine gegebene Funktion f das Integral

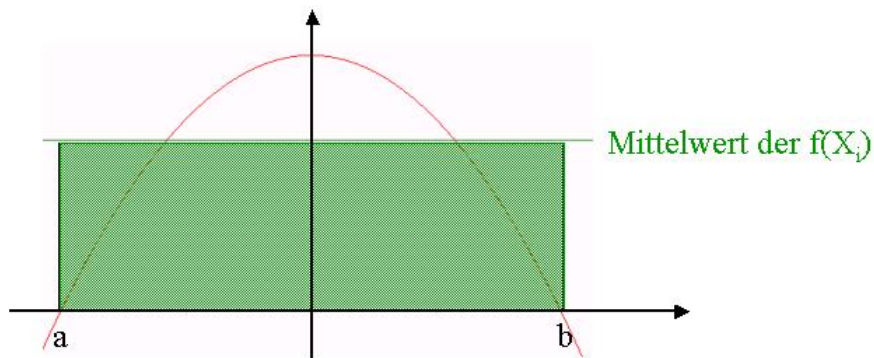
$$\int_a^b f(x) dx.$$

Dazu erzeugen wir N gleichverteilte Zufallszahlen $X_1, \dots, X_N$ (mit „sehr großem“ N) aus dem Intervall $[a, b]$ und bilden den Mittelwert über die sich daraus ergebenden Funktionswerte:

$$\frac{1}{N} \sum_{i=1}^N f(X_i).$$

Nun können wir das Integral $\int_a^b f(x)$ im Falle $f(a) = 0$, $f(b) = 0$ und $f(x) \geq 0$ für alle $x \in [a, b]$ aber anschaulich als den Flächeninhalt interpretieren, den der Graph von f mit der x -Achse einschließt. Diesen müsste man doch eigentlich dadurch annähern können, dass man die Länge des Intervalls, also $b - a$, mit dem Mittelwert der Funktionswerte multipliziert, also müsste anschaulich Folgendes gelten:

$$\int_a^b f(x) dx \approx (b - a) \cdot \frac{1}{N} \sum_{i=1}^N f(X_i).$$



Monte-Carlo-Simulation zur Berechnung von Integralen

Und tatsächlich gilt diese Formel, auch im allgemeineren Fall! Wir haben das oben im Programm bereits verwendet.

Kapitel 4

Simulation eines Erwartungswertes

Simulation 3: Die Entenjagd

Zehn Jäger, lauter perfekte Schützen (d.h. jedes anvisierte Ziel wird auch getroffen), lauern vor einem Fels auf Enten. Bald landen auf dem Fels zehn Enten. Die Jäger können nur einmal schießen und sie können nicht ausmachen, wer auf welche Ente schießt. Daher schießen sie gleichzeitig, aber jeder wählt sein Opfer zufällig aus. Wie viele Enten überleben im Durchschnitt, wenn dieses Experiment oft wiederholt wird?

Gruppenaufgabe:

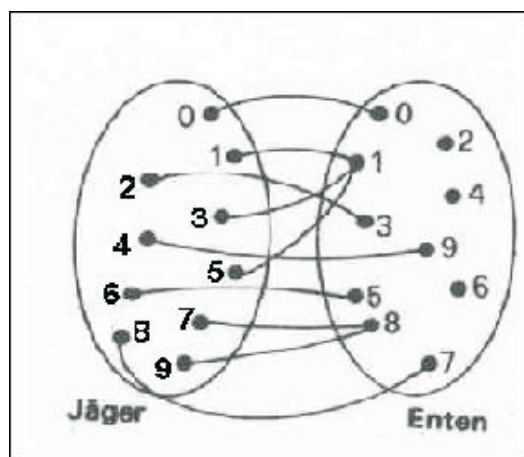
Bevor du weiterliest: Gib erst einmal eine Schätzung ab! Was vermutest du? Wie viele Enten überleben im Durchschnitt? Jetzt überlege dir: Wie könnte man dieses Experiment mit dem Computer simulieren? Hast du Ideen? Diskutiert in der Gruppe darüber, wie man das anstellen könnte.

Hier sieht man den Vorteil einer stochastischen Simulation: Ein echtes Experiment ist nicht möglich, da man keine zehn Schützen finden wird, die immer treffen und wir natürlich außerdem nicht tatsächlich auf lebende Enten schießen wollen, nur um Wahrscheinlichkeiten ermitteln zu wollen. Ei-

ne andere, tierfreundlichere Alternative wäre eine explizite mathematisch-analytische Berechnung des Durchschnittswertes. Das geht auch, aber man benötigt dafür Begriffe aus der Wahrscheinlichkeitsrechnung (den **Erwartungswert**), die ihr vermutlich noch nicht kennt. Aber auch wenn man diese Begriffe kennt, muss man erst einmal den richtigen „Kniff“ finden, wie man diese Aufgabe angeht. Wir werden uns das zwar unten gemeinsam überlegen, aber man kann nicht unbedingt jedes Mal erhoffen, selbst auf diese Ideen zu kommen (und es gibt natürlich noch viel kompliziertere Beispiele, wo man gar keine Chance hat, den Erwartungswert explizit zu ermitteln). Daher ist hier eine stochastische Simulation sehr hilfreich: Man kann mit Hilfe der Simulation einen Näherungswert für den Erwartungswert bestimmen und sich dann damit eine Idee für einen mathematisch-analytischen Nachweis besorgen. Auch umgekehrt kann die Simulation nützlich sein: Hat man den theoretischen Wert für die erwartete Anzahl der überlebenden Enten berechnet und ist sich nicht sicher, dass das Ergebnis auch wirklich richtig ist, so kann man es mit Hilfe der stochastischen Simulation überprüfen.

Das Gute ist: Obwohl die exakte Lösung nicht einfach zu finden ist, lässt sich das Experiment bequem und einfach simulieren, also „nachspielen“:

Man erzeugt ganz einfach 10 Zufallsziffern von 0 bis 9. Die fehlenden Ziffern sind die Nummern der überlebenden Enten, wobei wir bei 0 anfangen zu zählen. Die folgende Abbildung zeigt ein vom Ziffernblock 0131915878 gesteuertes Experiment:



Veranschaulichung des Ziffernblocks 0131915878 bei der Entenjagd

Die Schützen 0, 1, ..., 9 sind als Block 0123456789 kodiert. Der Ziffernblock 0131915878 kodiert die getroffenen Enten:

- Schütze 0 trifft Ente 0,
- Schütze 1 trifft Ente 1,
- Schütze 2 trifft Ente 3,
- Schütze 3 trifft auch Ente 1 usw.

Aufgabe 3:

Nehmen wir an, wir haben das Experiment simulativ durchgeführt und nach 5 Simulationsschritten die folgenden Blöcke erzeugt:

9416956090 ,
7506682320 ,
4628819077 ,
7592050242 ,
4233819129 .

Wie viele Enten haben in diesen 5 Simulationsschritten überlebt? Wie lautet somit der Mittelwert an überlebenden Enten?

Jetzt können wir die Simulation starten:

Gruppenaufgabe:

Schau dir das Excel-Programm zur Entenjagd an und versuche nachzuvollziehen, wie es arbeitet. Führe die Simulation nun für verschiedene Anzahlen durch und notiere dir die Werte.

Der theoretische Mittelwert (genauer: der Erwartungswert) für die Anzahl der überlebenden Enten beträgt übrigens:

$$m = 10 \cdot \left(\frac{9}{10}\right)^{10} \approx 3,4868.$$

Wir wollen nun, aber nur in Ansätzen, darauf eingehen, wie man genau darauf kommt: Nun, nehmen wir einmal an, wir hätten 10 Enten und einen Jäger. Dann würde natürlich immer genau eine Ente getroffen. Der Erwartungswert der getroffenen Enten wäre gleich 1. Nehmen wir nun an, wir hätten zwei Jäger. Dann gäbe es zwei Fälle:

Fall 1: Beide Jäger schießen auf die gleiche Ente.

Dies geschieht mit einer Wahrscheinlichkeit von $\frac{1}{10}$. In $\frac{1}{10}$ der Fälle wird dann also eine Ente getroffen.

Fall 2: Beide Jäger schießen auf verschiedene Enten.

Dies geschieht mit einer Wahrscheinlichkeit von $\frac{9}{10}$. In $\frac{9}{10}$ der Fälle werden also 2 Enten getroffen.

Führt man den Versuch also 10 mal durch, so wird in einem Fall 1 Ente und in 9 Fällen 2 Enten getroffen. Umgekehrt überleben also in einem Fall 9 Enten und in 9 Fällen 8 Enten. Daher ist die erwartete Anzahl der überlebenden Enten gleich:

$$\frac{1 \cdot 9 + 9 \cdot 8}{10} = \frac{81}{10}.$$

Es werden also in 10 Versuchen 81 Enten überleben, also statistisch 8,1 Enten pro Versuch.

Dieses Ergebnis kann man auch anders einsehen, und zwar diesmal aus Sicht der Enten: Schauen wir uns mal die erste Ente an: Wenn sie, rein hypothetisch, unendlich viele Leben hätte und zwei Jäger würde immer und immer wieder auf die 10 Enten schießen - wie oft würde sie dann überleben? Sie überlebt ja genau dann, wenn der erste Jäger nicht auf sie schießt (die Wahrscheinlichkeit dafür ist $\frac{9}{10}$) und wenn der zweite Jäger ebenfalls nicht auf sie

schießt (die Wahrscheinlichkeit dafür ist ebenfalls $\frac{9}{10}$). Die Wahrscheinlichkeit, dass unsere arme kleine erste Ente überlebt, ist also:

$$\frac{9}{10} \cdot \frac{9}{10} = \frac{81}{100}.$$

Sie wird demnach bei 100 Versuchen 81-mal überleben (rein hypothetisch natürlich).

Gleiches gilt auch für die zweite Ente. Sie wird erwartungsgemäß ebenfalls in 81 von 100 Versuchen überleben. Dies gilt natürlich für jede der zehn Enten. Daher werden in 100 Versuchen durchschnittlich

$$\underbrace{81 + 81 + \dots + 81}_{10\text{-mal}} = 10 \cdot 81 = 810$$

Enten überleben. Wiederum haben wir also eingesehen, dass pro Versuch erwartungsgemäß $\frac{810}{100} = 8,1$ Enten überleben werden.

Der Schritt auf 10 Enten und 10 Jäger ist jetzt nicht mehr schwierig! Wie groß ist die Wahrscheinlichkeit, dass die erste Ente überlebt? Sie überlebt genau dann, wenn sie von keinem der 10 Jäger getroffen wird. Die Wahrscheinlichkeit, dass sie von einem Jäger nicht getroffen wird, ist gleich der Wahrscheinlichkeit, dass der Jäger auf eine der anderen neun Enten schießt, also gleich $\frac{9}{10}$. Die Wahrscheinlichkeit, dass **alle** 10 Jäger auf eine andere Ente schießen (und unsere Ente somit überlebt) ist gleich

$$\left(\frac{9}{10}\right)^{10} = \underbrace{\frac{9}{10} \cdot \frac{9}{10} \cdot \dots \cdot \frac{9}{10}}_{10\text{-mal}}.$$

In 9^{10} von 10^{10} Fällen wird die Ente also überleben. Nun haben wir aber zehn Enten, d.h. in 10^{10} „Versuchen“ überleben erwartungsgemäß $10 \cdot 9^{10}$ Enten. Pro Versuch werden also durchschnittlich

$$\frac{10 \cdot 9^{10}}{10^{10}} = 10 \cdot \left(\frac{9}{10}\right)^{10} \approx 3,4868$$

Enten überleben.

Kapitel 5

Eine Kuriosität?

Simulation 4: Die Würfelschlange

Das folgende Experiment ist aus der Ausstellung „Mathematik zum Anfassen“ des Gießener Mathematikmuseums „Mathematikum“, geleitet von Herrn Professor Beutelspacher, bekannt:

Einige 6-seitige Würfel (sagen wir 60 Stück) werden zufällig in eine Reihe gelegt. Zuerst wählt man einen beliebigen der ersten sechs Würfel aus und zählt dann so viele Würfel weiter, wie der gewählte Würfel anzeigt. Nun schaut man sich den Würfel, auf dem man gelandet ist, an und zählt wieder so viele Würfel weiter, wie der Würfel anzeigt, usw. Wenn man dann am Ende der Schlange mehr Würfel abzählen müsste als vorhanden sind, hört man auf. Ist man zum Beispiel am vorletzten Würfel und zeigt dieser eine 2 oder mehr Augen, so hört man auf. Alle Würfel, die jetzt noch übrig sind, legt man beiseite. Anschließend fängt man mit einem anderen der ersten sechs Würfel an. Und dann mit noch einem weiteren...

Gruppenaufgabe:

Der Betreuer hat euch genügend Würfel mitgebracht, mit denen ihr das Experiment real durchführen könnt. Was fällt euch auf?

Der Witz ist folgender: Wiederholt man diese Prozedur mit einem anderen Anfangswürfel, so wird man fast immer beim gleichen Würfel der Schlange landen, bei dem man mit dem ersten Würfel als Startwürfel gelandet ist.

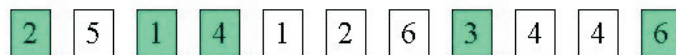
Gruppenaufgabe:

Führt jetzt die Simulation ein paar Mal durch, um diese Beobachtung zu bestätigen. Kannst du dir eine Erklärung für dieses Phänomen vorstellen? Diskutiert diese Frage untereinander.

Erklärung:

Wenn man alle erreichten Würfel des ersten Durchlaufs markiert, erhält man einen Pfad von Würfeln, der zu dem letzten erreichten Würfel führt. Sobald man erst einmal einen der Würfel dieses Pfades erreicht hat, wird man in jedem Fall am Schluss wieder bei dem gleichen Würfel landen.

Dazu ein kleines Beispiel mit 11 Würfeln:



Die türkis markierten Würfel gehören zum Pfad mit dem ersten Würfel als Startwürfel. Damit gehören auch der dritte und vierte Würfel zum Pfad. Startet man vom fünften oder sechsten Würfel, landet man ebenfalls auf dem Pfad (nämlich beim achten Würfel). Nur der zweite Würfel hat einen anderen Endwürfel (nämlich den siebten Würfel, danach geht es nicht mehr weiter).

Man sieht also, dass man **nicht immer automatisch** beim gleichen Würfel endet, dass es aber doch wahrscheinlich erscheint, dass dies so ist. Interessant ist demnach die Frage, wie groß die Wahrscheinlichkeit ist, den Pfad zu erreichen, den der erste Würfel vorgibt, oder wie groß umgekehrt die Wahrscheinlichkeit ist den Pfad bis zum Ende zu verpassen. Die exakte Wahrscheinlichkeit ist schwierig auszurechnen (und mir persönlich nicht gelungen). Man kann allerdings eine Abschätzung durchführen:

Man beginnt mit einem der ersten sechs Würfel, die nicht bereits zu demjenigen Pfad gehören, den der erste Würfel vorgibt. Die Chance, dass die darauf abgebildete Augenzahl zu keinem Pfadwürfel führt, ist höchstens $\frac{5}{6} \approx 0,93$, denn mindestens einer der nächsten sechs Würfel muss ja markiert sein, also zum Pfad gehören. Die Chance danach noch einmal keinen zu treffen, ist wieder höchstens $\frac{5}{6}$, also ist die Wahrscheinlichkeit, nach zwei Schritten den Pfad zu verfehlen, höchstens $\frac{5}{6} \cdot \frac{5}{6} \approx 0,69$. Damit ergibt sich eine Gesamtwahrscheinlichkeit von **höchstens** $\left(\frac{5}{6}\right)^n$ am Ende den Pfad **verfehlt** zu haben. Die Chance am selben Würfel zu enden, bei dem man mit dem ersten Würfel als Startwürfel endet, ist also **mindestens** gleich $1 - \left(\frac{5}{6}\right)^n$, wobei n die Anzahl der Schritte ist, die man machen muss um zum Ende der Schlange zu gelangen. Im „schlimmsten“ Fall sind dies (falls N , die Anzahl der vorhandenen Würfel, durch 6 teilbar ist) $\frac{N}{6}$ Schritte. (Dieser „schlimmste“ Fall tritt zum Beispiel dann ein, wenn der ausgewählte Startwürfel der sechste ist und die danach abgezählten Würfel alle eine 6 zeigen.) Falls also die Anzahl N der Würfel durch 6 teilbar ist, dann ist die Wahrscheinlichkeit auf dem Pfad, den der erste Würfel vorgibt, zu landen, mindestens gleich:

$$1 - \left(\frac{5}{6}\right)^{\frac{N}{6}}.$$

Wenn man also 60 Würfel hat, dann ist die Wahrscheinlichkeit am selben Würfel zu enden, bei dem man mit dem ersten Würfel als Startwürfel endet, mindestens gleich

$$1 - \left(\frac{5}{6}\right)^{10} \approx 0,84.$$

Dies erscheint nicht so hoch, wie es die Simulation erwarten lässt, aber man muss natürlich bedenken, dass wir hier nur eine **Abschätzung** vorgenommen haben und die theoretisch exakte Wahrscheinlichkeit für das Erreichen des Pfades sehr viel höher ist. Um dafür eine Annäherung zu bekommen, könnte man jetzt die Simulation sehr häufig durchlaufen lassen, um über die relative Häufigkeit eine Annäherung der theoretischen Wahrscheinlichkeit zu bekommen.

Rückwärtsbetrachtung:

Man kann das Feld auch von hinten aufrollen um herauszufinden, welche Würfel zum Pfad gehören können und welche der 6 möglichen Startwürfel

beim Würfel enden, bei dem man mit dem ersten Würfel als Startwürfel endet. Dazu betrachtet man den Endwürfel des vom ersten Würfel vorgegebenen Pfades. Dann überlegt man sich, welche vorhergehenden Würfel zu ihm führen und markiert diese. Mit diesen markierten Würfeln stellt man dieselben Überlegungen an, bis man am Anfang angelangt ist.



Um zum letzten Würfel zu kommen, muss man vorher bei einem der markierten Würfel gelandet sein.

Zusatzaufgabe:

Schau dir das Excel-Programm zur Würfelschlange jetzt noch einmal ganz in Ruhe an, auch den dahinterliegenden Excel-Code. Versuche nachzuvollziehen, wie es programmiert wurde.

Kapitel 6

Bedingte Wahrscheinlichkeiten

6.1 Simulation 5: Das Ziegenproblem

In der amerikanischen Spielshow „Let’s make a deal“ ist als Hauptpreis ein Auto ausgesetzt. Der Kandidat hat drei Türen zur Auswahl, zwischen denen er sich entscheiden muss. Hinter einer Tür befindet sich das Auto, hinter den beiden anderen jeweils eine Ziege. Der Quizmaster weiß natürlich, hinter welcher Tür sich das Auto befindet und öffnet nach der Wahl des Kandidaten eine Tür, hinter der sich eine Ziege befindet. Der Kandidat hat nun zwei Möglichkeiten: Entweder er bleibt bei seiner ursprünglichen Wahl oder aber er wählt die dritte, verbliebene Tür. Er erhält auf jeden Fall den Preis hinter der zuletzt von ihm gewählten Tür.



die amerikanische Spielshow „Let’s make a deal“

Gruppenaufgabe:

Spielt das Spiel jetzt ein paar Mal am Computer. Was fällt euch auf? Was meint ihr, auch aufgrund eurer Spielerfahrung: Ist es besser für den Kandidaten zu wechseln oder sollte er besser bei seiner ersten Wahl bleiben? Oder ist es völlig egal, ob er sich umentscheidet oder nicht, weil die Wahrscheinlichkeit auf den Autogewinn eh in beiden Fällen gleich hoch ist? Diskutiert ruhig darüber! Zumeist sind Diskussionen über dieses Problem recht emotional. ;-)

In der Kolumne „Ask Marilyn“ des amerikanischen Wochenmagazins „Parade“ erklärte die Journalistin Marilyn vos Savant ¹ den Leserinnen und Lesern, dass ein Wechsel des Kandidaten zu Tür 2 dessen Chancen im Vergleich zum Festhalten an Tür 1 verdoppeln würden.

Das war die Geburtsstunde des Ziegenproblems im Jahre 1991, denn die Antwort von Marilyn vos Savant führte zu einer Flut von Leserbriefen gegenteiliger Meinung. Etwa 90 Prozent der Zuschriften, darunter zahlreiche mathematisch hochgebildete Leute, waren der Meinung es sei egal, ob man wechsele oder nicht. Das Argument lautete dabei stets so: Die Gewinnchancen hätten sich durch den Hinweis des Moderators von $p = \frac{1}{3}$ auf $p = \frac{1}{2}$ erhöht, da jetzt „zwei gleichwahrscheinliche Türen übrig“ seien. Marilyn vos Savant blieb jedoch bei ihrer Empfehlung und argumentierte wie folgt:

Die Wahrscheinlichkeit, dass sich das Auto hinter der Tür 1 befindet, ist $\frac{1}{3}$. Die Wahrscheinlichkeit, dass es sich hinter **einer der beiden anderen** Türen befindet, ist $\frac{2}{3}$. Der Moderator öffnet nun eine der beiden Türen, aber **nicht zufällig**, denn er weiß ja, wo das Auto steht. Die Wahrscheinlichkeit, dass es eine der beiden zu Beginn nicht gewählten Türen war, bleibt aber $\frac{2}{3}$. Es steht nun nur halt für den Außenstehenden fest, welche der beiden Türen die Wahrscheinlichkeit von $\frac{2}{3}$ trägt.

¹Marilyn vos Savant wird vom Guinness Buch der Rekorde in der Rubrik „Ruhmeshalle“ unter „Höchster IQ“ aufgeführt. Ihre Kolumne für scharfe Denker und solche, die es werden wollen, ist seit Jahren ein Riesenerfolg in den USA. Die Autorin ist Mitglied im Vorstand des National Council on Economic Education und wirkt mit im Beirat des Verbands für hochbegabte Kinder. Sie lebt mit ihrem Mann Robert Jarvik, dem Erfinder des künstlichen Herzens, in New York City.

Diese Argumentation ist nicht immer für jeden einsichtig, auch wenn sie richtig ist. Viel deutlicher versteht man das Problem, wenn man sich überlegt, mit welcher Strategie man wann genau gewinnt.

Der Standhafte gewinnt genau dann das Auto, wenn sich dieses hinter der ursprünglich gewählten Tür befindet, und die Wahrscheinlichkeit dafür ist $\frac{1}{3}$. Ein Wechsler gewinnt hingegen das Auto genau dann, wenn er sich zuerst für eine „Ziegentür“ entschieden hat (die Wahrscheinlichkeit hierfür ist $p = \frac{2}{3}$) und in diesem Fall beim Wechseln ja durch das Öffnen der zweiten „Ziegentür“ durch den Moderator automatisch auf die „Autotür“ geführt wird.

Das Problem ist also, dass die Chancen nach dem Öffnen der Tür durch den Moderator eben nicht 50:50 sind, sondern davon abhängen, ob man sich zu Beginn für die richtige Tür oder aber die falsche Tür entschieden hat.

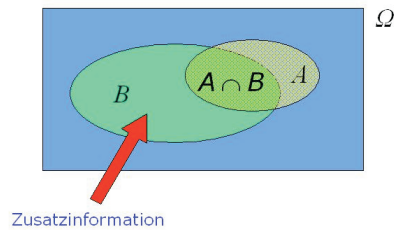
Wir wollen dieses Problem nun mathematisch lösen. Dazu benötigen wir zunächst den Begriff der **bedingten Wahrscheinlichkeit**:

Definition 2 (bedingte Wahrscheinlichkeiten: $P(A|B)$)

Die bedingte Wahrscheinlichkeit $P(A|B)$ des Ereignisses A bei gegebenem Ereignis B mit $P(B) > 0$ ist definiert als Quotient

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

Man kann $P(A|B)$ als die Wahrscheinlichkeit dafür ansehen, dass das Ereignis A eintritt, wenn man schon weiß, dass das Ereignis B eingetroffen ist, also als Wahrscheinlichkeit von A unter der **Zusatzinformation** B :



bedingte Wahrscheinlichkeiten als Zusatzinformationen

Beispiel:

Wir betrachten wieder ein Würfelexperiment, also den Ereignisraum $\Omega = \{1, 2, 3, 4, 5, 6\}$ mit den folgenden Elementarwahrscheinlichkeiten:

$$P(\{\omega\}) = \frac{1}{6},$$

also:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{6}.$$

Weiterhin sei B das Ereignis, dass man eine gerade Zahl würfelt, also: $B = \{2, 4, 6\}$, und A sei das Ereignis, dass man eine 6 erhalten hat, also: $A = \{6\}$.

Wie groß ist die Wahrscheinlichkeit, dass man eine 6 gewürfelt hat, wenn schon bekannt ist, dass man eine gerade Zahl gewürfelt hat?

Es gilt:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{6\})}{P(\{2, 4, 6\})} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}.$$

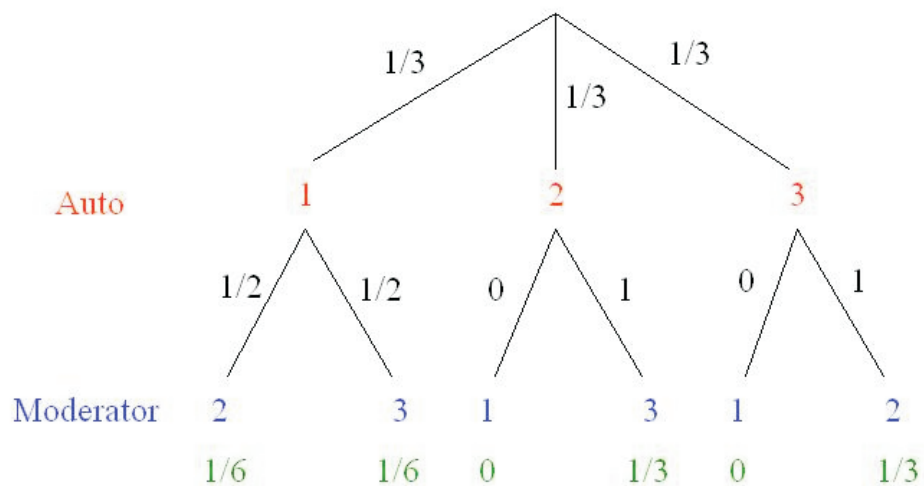
Anschaulich klar: Man hat drei gerade Zahlen, darunter ist genau eine die 6.

Aufgabe 4:

- (a) Wie groß ist die Wahrscheinlichkeit dafür, dass man eine Zahl gewürfelt hat, die kleiner als 4 ist, wenn bereits bekannt ist, dass man eine Primzahl gewürfelt hat?
- (b) Wie groß ist die Wahrscheinlichkeit dafür, dass man weder eine 3 noch eine 5 gewürfelt hat unter der Voraussetzung, dass man keine 4 gewürfelt hat?

Mathematische Lösung des Ziegenproblems

Entscheidet sich der Kandidat zunächst für die Tür Nummer 1, dann ergibt sich das folgende Baumdiagramm:



Baumdiagramm für das Ziegenproblem

Daraus berechnen wir jetzt die bedingten Wahrscheinlichkeiten unter der Voraussetzung, dass der Moderator die Tür 3 öffnet und sich dahinter eine Ziege verbirgt. (Das Prinzip in dem Fall, wo der Moderator die Tür 2 öffnet und sich dahinter die Ziege befindet, ist natürlich das gleiche. Wir beschränken uns daher auf den Fall, bei dem der Moderator die Tür 3 öffnet.) Ist es dann wahrscheinlicher, dass sich das Auto hinter der Tür 2 oder hinter der Tür 1 befindet? Oder ist es gleichwahrscheinlich? Ist es dann besser zu wechseln oder bei seiner alten Wahl zu bleiben?

Welcher Wert ist also höher:

$P(\text{das Auto befindet sich hinter Tür 2} \mid \text{der Moderator öffnet Tür 3})$

oder

$P(\text{das Auto befindet sich hinter Tür 1} \mid \text{der Moderator öffnet Tür 3})$?

$$\begin{aligned}
 & P(\text{das Auto befindet sich hinter Tür 2} \mid \text{der Moderator öffnet Tür 3}) \\
 = & \frac{P(\text{das Auto befindet sich hinter Tür 2 und der Moderator öffnet Tür 3})}{P(\text{der Moderator öffnet Tür 3})} \\
 = & \frac{\frac{1}{3}}{\frac{1}{6} + \frac{1}{3}} \\
 = & \frac{\frac{1}{3}}{\frac{1}{2}} \\
 = & \frac{2}{3}
 \end{aligned}$$

und

$$\begin{aligned}
& P(\text{das Auto befindet sich hinter Tür 1} \mid \text{der Moderator öffnet Tür 3}) \\
&= \frac{P(\text{das Auto befindet sich hinter Tür 1 und der Moderator öffnet Tür 3})}{P(\text{der Moderator öffnet Tür 3})} \\
&= \frac{\frac{1}{6}}{\frac{1}{6} + \frac{1}{3}} \\
&= \frac{\frac{1}{6}}{\frac{1}{2}} \\
&= \frac{1}{3}.
\end{aligned}$$

Die bedingte Wahrscheinlichkeit (unter der Bedingung, dass der Moderator die Tür 3 geöffnet hat) dafür, dass sich das Auto hinter der verbliebenen dritten Tür befindet, ist also mit $p = \frac{2}{3}$ doppelt so hoch wie die bedingte Wahrscheinlichkeit dafür, dass sich das Auto hinter der ersten Türwahl verbirgt.

Daher ist es besser zu wechseln. Ein Wechsel verdoppelt die Gewinnchancen!

Bemerkung:

Natürlich muss man sich klarmachen, dass wahrscheinlichkeitstheoretische und statistische Aussagen natürlich vor allem für langfristige Prognosen und vielfache Wiederholungen relevant sind. Im Einzelfall, d.h. wenn man wirklich an der Show teilnimmt, sollte man zwar rein rational auch die Tür wechseln (denn es verdoppelt auf jeden Fall die Gewinnchance!), aber dort spielen dann vermutlich auch psychologische Faktoren eine große Rolle. Wer will sich hinterher schon den Vorwurf machen zur falschen Tür gewechselt zu sein, wenn man anfangs die richtige Tür gewählt hat?

Aufgabe 5:

Wir spielen das Ziegenspiel jetzt einmal mit vier Türen. Hinter einer der vier Türen befindet sich ein Auto, hinter drei der vier Türen eine Ziege. Nehmen wir weiter an, du hast dich für die erste Tür entschieden und der Moderator öffnet die zweite Tür. Ist es dann besser zur dritten (oder vierten) Tür zu wechseln oder bleibst du besser bei deiner ersten Wahl? Oder ist es völlig egal, da die Gewinnchancen eh die gleichen sind? Male dir wie oben ein Baumdiagramm auf und berechne die bedingten Wahrscheinlichkeiten. Wie sieht das Ganze wohl mit 100 Türen und einer Ziege aus?

Gruppenaufgabe:

Wir versuchen uns nun zusammen am Smartboard einen Algorithmus für die Programmierung der Simulation zu überlegen.

Bemerkung:

Man kann das Problem auch ganz allgemein für n Türen und k Ziegen betrachten und lösen. Das aber würde hier zu weit führen. Vielleicht aber hast du ja Lust, im Internet nach Lösungen dieses allgemeinen Problems zu suchen und diese zu verstehen!