15 Diskrete Wahrscheinlichkeitsräume

15.1 Grundbegriffe Ω sei eine endliche oder abzählbar unendliche Menge. Eine Abbildung $P : \Omega \to [0,1]$ heißt *Wahrscheinlichkeitsmaß*, wenn $\sum_{\omega\in\Omega} P(\omega) = 1$. In diesem Fall wird (Ω, P) als *Wahrscheinlichkeitsraum* bezeichnet, der aus *elementaren Ereignissen* $\omega \in \Omega$ besteht.

Anschaulich besteht ein Wahrscheinlichkeitsraum aus den möglichen Ausgängen eines wiederholbaren Zufallsexperiments, wie dem Werfen eines Würfels ($\Omega = \{1, \dots, 6\}$), dem Ziehen einer Karte aus einem Kartenspiel ($\Omega = \{1, \dots, 52\}$), oder dem Werfen einer Münze ($\Omega = \{K, Z\}$). $P(\omega)$ gibt dann die relative Häufigkeit an, mit der ω eintritt. $P(\omega) = 1$ bedeutet demnach, dass ω immer eintritt, wegen $\sum P(\omega) = 1$ treten dann alle anderen elementaren Ereignisse niemals ein. In den oben genannten Beispielen kann man davon ausgehen, dass alle elementaren Ereignisse *gleichverteilt* sind und demnach $P(\omega) = 1/|\Omega|$ gilt, wobei $|\Omega|$ die Anzahl der Elemente von Ω bezeichnet.

$A \subset \Omega$ heißt *Ereignis*, die *Wahrscheinlichkeit* von A ist dann definiert durch

$$P(A) = \sum_{\omega\in A} P(\omega).$$

Aus dieser Definition folgt für $A, B, A_i \subset \Omega$

$$P(A^c) = 1 - P(A), \quad \text{insbesondere } P(\emptyset) = 0,\ P(\Omega) = 1$$

$$A \subset B \ \Rightarrow\ P(A) \le P(B)$$

$$P(A \setminus B) = P(A) - P(A \cap B)$$

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

$$P(\cup_{i=1}^n A_i) \le \sum_{i=1}^n P(A_i) \quad \text{mit Gleichheit bei disjunkten } A_i$$

Im Fall einer Gleichverteilung bedeutet die Definition

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{„Zahl der günstigen Fälle”}}{\text{„Zahl der möglichen Fälle”}}$$

Wir behandeln nun die vier grundsätzlichen Abzählprobleme bei Gleichverteilung. In einer Urne liegen N Kugeln, die wir uns von 1 bis N nummeriert denken. Es werden n Kugeln sukzessive gezogen, was *Stichprobe* genannt wird. Wie viele solcher Stichproben es gibt, hängt von der Art der Ziehung ab und davon, ob die Reihenfolge der gezogenen Kugeln berücksichtigt wird.

I Stichproben in Reihenfolge mit Rücklegen Nach jedem Ziehen wird die Kugel wieder zurückgelegt, unterschiedliche Reihenfolgen werden mitgezählt. Dies ist äquivalent dazu, aus n Urnen jeweils eine Kugel zu ziehen. Die Zahl der Möglichkeiten ist demnach N^n.

II Stichproben in Reihenfolge ohne Rücklegen Für die erste Kugel gibt es N Möglichkeiten, für die nächste $N-1$. Für $n \le N$ gibt es daher $N(N-1)\dots(N-n+1)$ Möglichkeiten, für $n > N$ keine.

III Stichproben ohne Reihenfolge ohne Rücklegen In diesem Fall stimmt die Zahl der Möglichkeiten mit der Anzahl der n-elementigen Teilmengen von $\{1, \dots, N\}$ überein, das sind $\binom{N}{n}$.

IV Stichproben ohne Reihenfolge mit Rücklegen Wir normieren die Stichprobe, indem wir die gezogenen Kugeln nach Größe ordnen. Demnach ist die Zahl der Elemente der Menge

$$M = \big\{(\omega_1, \omega_2, \dots, \omega_n) : \omega_1 \leq \omega_2 \leq \dots \leq \omega_n,\, \omega_i \in \{1, \dots, N\}\big\}$$

zu bestimmen. Wir addieren die i-te Komponente eines Vektors in M mit $i-1$ und erhalten so einen Vektor der Menge

$$M' = \big\{(\omega_1', \omega_2', \dots, \omega_n') : \omega_1' < \omega_2' < \dots < \omega_n',\, \omega_i \in \{1, \dots, N+n-1\}\big\}$$

Da aber auch umgekehrt aus jedem Vektor in M', wenn man von seiner i-ten Komponente $i-1$ abzieht, ein Element von M entsteht, sind beide Mengen gleich groß. Damit gibt es $\binom{N+n-1}{n}$ Stichproben ohne Reihenfolge mit Rücklegen.

Beispiele 15.1 (i) Es werden vier nicht unterscheidbare Würfel gleichzeitig geworfen. Wie groß ist die Wahrscheinlichkeit p, dass die vier erscheinenden Augenzahlen verschieden sind?

Der zugrunde liegende Wahrscheinlichkeitsraum ist die Zahl aller möglichen Würfe, also 6^4. Um die Zahl der günstigen Fälle zu bestimmen, macht man sich klar, dass es für den ersten Würfel 6 Möglichkeiten gibt, für den zweiten 5 usw. Es liegt also Typ II vor und wir erhalten

$$p = \frac{6 \cdot 5 \cdot 4 \cdot 3}{6^4} = \frac{5}{18}.$$

(ii) Wie groß ist die Wahrscheinlichkeit p, beim Zahlenlotto „6 aus 49“ genau vier Richtige zu haben?

Da die Reihenfolge der gezogenen Zahlen keine Rolle spielt, ist der Wahrscheinlichkeitsraum vom Typ III, seine Kardinalität daher $\binom{49}{6}$. Man hat genau vier richtige Zahlen, wenn vier Zahlen in der Menge der sechs gezogenen Zahlen liegen und zwei Zahlen außerhalb. Da es in beiden Fällen nicht auf die Reihenfolge ankommt, liegt wieder Typ III vor und wir erhalten

$$p = \frac{\binom{6}{4}\binom{43}{2}}{\binom{49}{6}} = 0,00096\dots$$

□

Das letzte Beispiel ist ein Spezialfall der *hypergeometrischen Verteilung*: Aus einer Urne mit S schwarzen und W weißen Kugeln werden $n \leq N = S + W$ Kugeln ohne Rücklegen gezogen. Die Wahrscheinlichkeit dafür, dass diese Stichprobe aus genau s schwarzen und $w = n - s$ weißen Kugeln besteht, beträgt dann

$$h(s; n, N, S) = \binom{S}{s}\binom{W}{w} / \binom{S+W}{n}.$$

Um dies einzusehen, argumentieren wir genauso wie im letzten Beispiel. Die Zahl der möglichen Stichproben (Typ III) ist $\binom{S+W}{n}$. Da wir s schwarze Kugeln erhalten wollen, haben wir dazu $\binom{S}{s}$ Möglichkeiten. Da die zu ziehenden weißen Kugeln von den gezogenen schwarzen Kugeln unabhängig sind, muss diese Zahl mit der Zahl der Möglichkeiten, $n - s$ weiße Kugeln zu ziehen, multipliziert werden.

Ist R eine Aussage über die Elemente eines Wahrscheinlichkeitsraumes, so ist $P(R)$ die Summe der $P(\omega)$, für die $R(\omega)$ wahr ist.

15.2 Bedingte Wahrscheinlichkeit und Unabhängigkeit Mit $P(A \mid B)$ bezeichnen wir die Wahrscheinlichkeit des Ereignisses A, wenn wir bereits wissen, dass B eingetreten ist. Als Beispiel betrachten wir das Werfen eines Würfels. A sei das Ereignis, dass eine 6 geworfen wurde, B das Ereignis, dass eine gerade Zahl geworfen wurde. Wenn also B bereits eingetroffen ist, bleiben nur nur noch die Möglichkeiten $2, 4, 6$, die Wahrscheinlichkeit für A ist dann $1/3$. Bei Gleichverteilung erhalten wir daher für die bedingte Wahrscheinlichkeit

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

Mit $P(A \cap B) = |A \cap B|/|\Omega|$ und $P(B) = |B|/|\Omega|$ ist dies äquivalent zu

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}. \tag{15.1}$$

Sei nun (Ω, P) ein allgemeiner Wahrscheinlichkeitsraum. Ist $P(B) > 0$, so definieren wir *die bedingte Wahrscheinlichkeit $P(A \mid B)$ von A bei gegebenem B* durch (15.1). Diese Definition kann leicht mit der anschaulichen Bedeutung von P als relativer Häufigkeit eines Zufallsexperiments erklärt werden. Wir brauchen nur die Fälle des Experiments zu betrachten, in denen das Ereignis B eintrifft. Dann liegt entweder der günstige Fall $A \cap B$ oder der ungünstige Fall $B \setminus A$ vor.

Lemma 15.2 *Sind $A_1, \ldots, A_k$ Ereignisse mit $P(A_1), \ldots, P(A_k) > 0$, so gilt*

$$P(A_1 \cap \ldots \cap A_k) = P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cap A_2) \cdot \ldots \cdot P(A_k \mid A_1 \cap A_2 \cap \ldots \cap A_{k-1})$$

Beweis: Wir verwenden Induktion über k. $k = 2$ ist gerade die Definition der bedingten Wahrscheinlichkeit. Für den Induktionsschritt schreiben wir

$$P(A_1 \cap \ldots \cap A_{k-1} \cap A_k) = P(A_1 \cap \ldots \cap A_{k-1}) \cdot P(A_k \mid A_1 \cap \ldots \cap A_{k-1}).$$

□

Beispiel 15.3 *Skat* wird mit einem deutschen Kartenspiel zu 32 Karten gespielt. Die drei Spieler erhalten jeweils 10 Karten, die restlichen 2 Karten werden in den *Skat* gelegt. Wie groß ist die Wahrscheinlichkeit, dass jeder Spieler genau ein As besitzt?

Wir nehmen an, dass Spieler 1 die ersten 10 Karten, Spieler 2 die zweiten 10 und Spieler 3 die dritten 10 Karten bekommt. Sei A_i das Ereignis, dass Spieler i genau ein As erhält. Nach dem letzten Lemma gilt dann

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cap A_2).$$

Es ist $P(A_1) = \binom{4}{1}\binom{28}{9}/\binom{32}{10}$ und $P(A_2 \mid A_1) = \binom{3}{1}\binom{19}{9}/\binom{22}{10}$, denn nachdem Spieler 1 die 10 Karten mit genau einem As bekommen hat, bleiben noch 22 Karten mit drei Assen übrig. Analog gilt $P(A_3 \mid A_1 \cap A_2) = \binom{2}{1}\binom{10}{9}/\binom{12}{10}$. □

Satz 15.4 (a) *Sei $P(B) > 0$. Durch $P_B(A) = P(A \mid B)$ ist ein Wahrscheinlichkeitsmaß auf Ω definiert.*

(b) *(Formel von der totalen Wahrscheinlichkeit) Sei $\{B_1, \ldots, B_k\}$ eine disjunkte Zerlegung von Ω. Dann gilt für jedes Ereignis A*

$$P(A) = \sum_i P(B_i) P(A \mid B_i),$$

wobei im Falle $P(B_i) = 0$ das Produkt $P(B_i)P(A|B_i)$ zu Null gesetzt wird.

(c) *(Formel von Bayes) Ist $P(A) > 0$ und gelten die Voraussetzungen von (ii), so*

$$P(B_j \mid A) = \frac{P(B_j)\, P(A \mid B_j)}{\sum_i P(B_i)P(A \mid B_i)}.$$

Beweis: (a) ist klar.

(b) folgt aus

$$P(A) = P\big(\cup_i (A \cap B_i)\big) = \sum_i P(A \cap B_i) = \sum_i P(B_i)P(A \mid B_i).$$

(c) folgt aus (b) wegen

$$P(B_j \mid A) = \frac{P(B_j \cap A)}{P(A)} = \frac{P(B_j)\, P(A \mid B_j)}{\sum_i P(B_i)P(A \mid B_i)}.$$

□

Beispiel 15.5 Eine Krankheit kommt bei $0,5\%$ der Bevölkerung vor. Es gibt einen Test für diese Krankheit, der 99% der Kranken identifiziert, aber auch bei 2% der Gesunden positiv ist. Wie groß ist die Wahrscheinlichkeit, dass ein positiv Getesteter tatsächlich krank ist?

Sei $\{1, \ldots, N\}$ die Menge der Bevölkerung, B_k sei die Menge der Kranken, B_g die der Gesunden. Damit gilt $|B_k| \approx 0,005N$, $|B_g| \approx 0,995N$. Ist ferner A die Menge der positiv getesteten Personen, so haben wir $|A \cap B_K| \approx 0,99|B_k|$ und $|A \cap B_g| \approx 0,02|B_g|$. Bei zufälliger Auswahl einer Person ergeben sich die Wahrscheinlichkeiten

$$P(B_k) = 0,005,\ P(B_g) = 0,995,\ P(A \cap B_k) = 0,99 \cdot 0,005,\ P(A \cap B_g) = 0,02 \cdot 0,995,$$

nach der Formel von Bayes daher

$$P(B_k \mid A) = \frac{P(A \cap B_k)}{P(A)} = \frac{0,99 \cdot 0,005}{0,99 \cdot 0,005 + 0,02 \cdot 0,995} = \frac{495}{2485} \approx 0,2.$$

Von den postiv getesteten Personen sind demnach nur 20% krank. □

Seien A, B zwei Ereignisse mit Wahrscheinlichkeiten $P(A), P(B) > 0$. Wir können $P(A)$ als eine Wette auf das Ereignis A ansehen. A und B definieren wir als unabhängig, wenn die Kenntnis von B an der Wahrscheinlichkeit für das Eintreffen von A nichts ändert, wenn also $P(A) = P(A \mid B)$. Anders ausgedrückt: Die Kenntnis von B hilft uns nicht weiter.

Der Begriff der Unabhängigkeit ist problematisch, weil er von der Mehrheit der Menschen nicht akzeptiert wird. Es ist aber eine empirische Tatsache, dass der zweite Wurf einer Münze oder einer Roulettekugel nicht vom ersten Wurf in irgendeiner Weise beeinflußt wird. Gegen diese Tatsache stehen Millionen Menschen, die glauben, dass aus den ausgespielten Zahlen beispielsweise im Lotto oder im Roulette auf die Zahlen der Zukunft geschlossen werden kann. Die gängigsten Einwände gegen die Unabhängigkeit von Ereignissen kann unter der Theorie subsummiert werden, dass jedes Ereignis ein raum-zeitliches Ereignisfeld aufbaut, das auch die anderen Ereignisse beeinflußt. Dem Einwand, das eine solche Einflußnahme empirisch nicht nachweisbar ist, kann mit Kritik am Versuchsaufbau oder dem Argument begegnet werden, dass die Einflußname äußerst schwach ist. Eine argumentative Auseinandersetzung mit diesem Standpunkt ist nicht möglich, weil die Wahrscheinlichkeit ein mathematisches Konstrukt ist, das keine Aussagen über die Tiefenstruktur der Welt erlaubt, zumal die Unabhängig von Ereignissen bei sehr kleinen Teilchen durchaus zweifelhaft ist.

Wir werfen zwei Würfel hintereinander, A sei das Ereignis, dass im ersten Wurf eine 1 oder 2 erscheint, B das Ereignis, dass im zweiten Wurf eine 6 kommt. Dann gilt

$$P(A) = \frac{1}{3}, \quad P(B) = \frac{1}{6}, \quad P(A \cap B) = \frac{1}{18}.$$

Damit ist $P(A \,|\, B) = \frac{1}{3} = P(A)$, was auch anschaulich klar ist, denn das Auftreten von B ändert nichts an der Wahrscheinlichkeit von A. Um eine korrekte mathematische Definition zu bekommen, die auch im Falle $P(B) = 0$ greift, bezeichnen wir zwei Ereignisse A und B als *unabhängig*, wenn

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

erfüllt ist, denn in diesem Fall gilt $P(A) = P(A \,|\, B)$ und $P(B) = P(B \,|\, A)$.

Beispiel 15.6 Wir werfen zwei Würfel hintereinander. A sei das Ereignis, dass eine gerade Augensumme gewürfelt wurde, B sei das Ereignis, das die zweite Augenzahl geradzahlig ist. Dann gilt

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{1}{2}, \quad P(A \cap B) = \frac{1}{4},$$

die beiden Ereignisse sind also unabhängig, obwohl B durchaus A beeinflußt. Der zugrunde liegende Mechanismus wird deutlicher, wenn wir die Wahrscheinlichkeit für das Auftreten einer geraden Zahl in beiden Würfen zu $2/5$ abändern. Dann gilt

$$P(A) = \big(\frac{2}{5}\big)^2 + \big(\frac{3}{5}\big)^2, \quad P(A \cap B) = \big(\frac{2}{5}\big)^2 \neq P(A)P(B).$$

A und B sind nun nicht mehr unabhängig. $\square$

Sei I eine beliebige Indexmenge und $\{A_i : i \in I\}$ eine Familie von Ereignissen. Die Familie heißt *unabhängig*, wenn für alle endlichen Indexmengen $\emptyset \neq J \subset I$ gilt

$$P\Big(\bigcap_{i \in J} A_i\Big) = \prod_{i \in J} P(A_i). \tag{15.2}$$

Beispiel 15.7 Der Sinn dieser Definition wird folgendermaßen deutlich. Wir werfen zweimal hintereinander eine Münze und betrachten die Ereignisse A=„beim ersten Wurf K“, B=„beim zweiten Wurf K“, C=„genau ein K“. Dann sind natürlich A und B unabhängig, aber, obwohl A und B das Ereignis C bestimmen, sind A und C ebenfalls unabhängig wegen $P(A) = 1/2$, $P(C) = 1/2$, $P(A \cap C) = 1/4$. Gleiches gilt für B und C. Andererseits ist $P(A \cap B \cap C) = 0$, aber $P(A)P(B)P(C) = 1/8$. Aus der paarweisen Unabhängigkeit folgt daher nicht notwendig die oben definierte Unabhängigkeit. $\square$

Satz 15.8 (a) *Ist $\{A_i : i \in I\}$ eine unabhängige Familie von Ereignissen und ist $P(A_k) = 0$ oder $P(A_k) = 1$ mit $k \notin I$, so ist auch $\{A_i : i \in I \cup \{k\}\}$ unabhängig.*

(b) *Ist $\{A_i : i \in I\}$ unabhängig und ist $B_i \in \{A_i, A_i^c, \emptyset, \Omega\}$ für jedes $i \in I$, so ist auch $\{B_i : i \in I\}$ unabhängig.*

(c) *Ist $I = \{1, \ldots, n\}$, so ist die endliche Familie von Ereignissen $\{A_i : i \in I\}$ genau dann unabhängig, wenn für jede Wahl von $B_i \in \{A_i, A_i^c\}$ gilt*

$$P(B_1 \cap \ldots \cap B_n) = P(B_1) \ldots P(B_n). \tag{15.3}$$

Beweis: (a) Gehört k zur Indexmenge J in (15.2), so ändert sich im Falle $P(B) = 1$ auf beiden Seiten der Definition nichts wegen $P(A \cap B) = P(A)$ für alle Mengen A. Im Falle $P(B) = 0$ steht hingegen auf beiden Seiten Null.

(b) Wegen (a) ist nur der Fall $B_i \in \{A_i, A_i^c\}$ näher zu untersuchen. Durch Induktion über m beweisen wir die Behauptung: Ist $J \subset I$ endlich und $|\{j \in J : B_j = A_j^c\}| \le m$, so gilt

$$P\Big(\bigcap_{j\in J} B_j\Big) = \prod_{j\in J} P(B_j). \tag{15.4}$$

Ist $m = 0$, so sind alle B_i gleich A_i und (15.4) folgt aus (15.2). Sei (15.4) für m bewiesen und J eine Indexmenge mit $|\{j \in J : B_j = A_j^c\}| = m + 1$. Wir können annehmen. dass $J = \{1, \ldots, n\}$ und $B_1 = A_1^c$. Für $j = 2, \ldots, n$ kann dann die Induktionsvoraussetzung verwendet werden. Es gilt daher

$$\begin{aligned} P\Big(\bigcap_{j=1}^n B_j\Big) &= P\Big(\bigcap_{j=2}^n B_j\Big) - P\Big(A_1 \cap \bigcap_{j=2}^n B_j\Big) \\ &= \prod_{j=2}^n P(B_j) - P(A_1)\cdot \prod_{j=2}^n P(B_j) = \prod_{j=1}^n P(B_j). \end{aligned}$$

(c) Die Notwendigkeit der Bedingung ist gerade (b). Auf (15.3) addieren wir die gleiche Bedingung für die Mengen $B_1^c, B_2, \ldots, B_n$ und erhalten

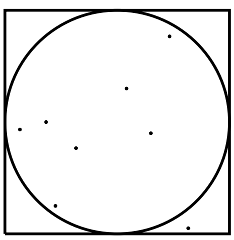
$$\begin{aligned} P\Big(\bigcap_{i=2}^n B_i\Big) &= P\Big(B_1^c \cap \bigcap_{i=2}^n B_i\Big) + P\Big(\bigcap_{i=1}^n B_i\Big) \\ &= P(B_1^c)\prod_{i=2}^n P(B_i) + \prod_{i=1}^n P(B_i) = \prod_{i=2}^n P(B_i). \end{aligned}$$

Auf diese Weise können wir die Bedingung auch für kleinere Indexmengen als $\{1, \ldots, n\}$ nachweisen.
$\square$

15.3 Die Monte-Carlo-Methode

Berechnung des Kugelvolumens Sei

$$K_n = \{x \in \mathbb{R}^n : \; x_1^2 + x_2^2 + \ldots + x_{n-1}^2 + x_n^2 < 1\}.$$



die Einheitskugel des n-dimensionalen Raums. In $n = 2$ Dimensionen sperren wir den Einheitskreis in das Quadrat $Q_1 = (-1, 1) \times (-1, 1)$ ein. Wir generieren Zufallsvektoren $(x_1, x_2) \in Q_1$, die auf Q_1 gleichverteilt sind, und zählen, wie viele davon in den Einheitskreis fallen. Das Verhältnis der Treffer zur Gesamtzahl der Versuche konvergiert dann gegen den gesuchten Flächeninhalt/4.

Bei N Versuchen erhalten wir für die Kugeln in den Raumdimensionen $n = 2, 3, 4, 5$:

n	$N = 10^1$	$N = 10^3$	$N = 10^5$	$N = 10^7$
2	3,59999	3,25200	3,14276	3,14222
3	4,80000	4,23199	4,18344	4,18810
4	3,20000	4,97599	4,90384	4,93469
5	9,60000	6,07999	5,28032	5,26119

$\mathrm{Vol}(K_2) = 3,14159\ldots$ liefert für die Fehler im Falle $n = 2$:

n	$K = 10$	$K = 10^3$	$K = 10^5$	$K = 10^7$
2	0.45841	0.11041	0.0117	0.0063

An diesem einfachen Beispiel sehen wir die Vor- und Nachteile der Monte-Carlo-Simulation:

1. Man braucht vom gestellten Problem keine Ahnung zu haben. In unserem Fall lässt sich das Kugelvolumen durch höhere Mathematik exakt bestimmen.

2. Daraus folgt: Auch sehr komplexe Probleme können mit einfachen Monte-Carlo-Prgrammen behandelt werden.

3. Das Verfahren ist sehr langsam, in der Regel geht der Fehler wie $\sqrt{N^{-1}}$.

4. Das Verfahren ist nicht sicher. Es gibt immer Ausreißer, die stark von den angegebenen $\sqrt{N^{-1}}$ abweichen.

Weiter benötigt man sehr viele Zufallszahlen. Woher nimmt man die?

Zufallszahlen Für die Bestimmung des Kugelvolumens reichen gleichverteilte Zufallszahlen über dem Intervall $[0, 1)$ aus. Durch eine Streckung oder Stauchung dieses Intervalls kann man Zufallszahlen auf jedem anderen Intervall bekommen. Gleichverteilt bedeutet, dass die Zufallszahl y mit Wahrscheinlichkeit $b - a$ im Intervall $(a, b) \subset [0, 1)$ liegt.

Da es sehr schwierig ist, „echte“ Zufallszahlen zu generieren, verwendet man durch eine Formel erzeugte „Pseudo-Zufallszahlen“. Im *Linearen Kongruenz-Generator* gibt man sich eine natürliche Zahl m vor sowie Zahlen $a, b \in \{0, 1, \dots, m-1\}$. Für eine weitere vorgegebene Zahl $x_0 \in \{0, 1, \dots, m-1\}$ als Startwert bestimmt man

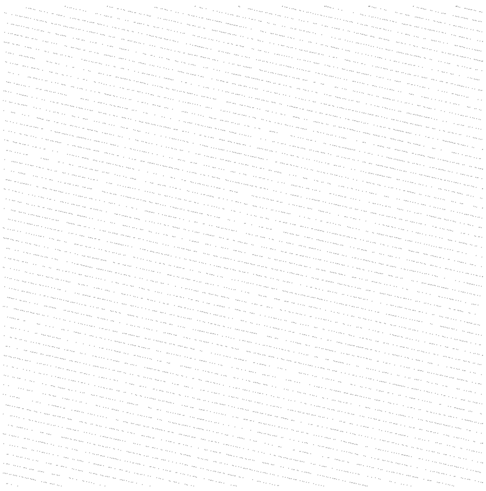
$$x_{i+1} = (ax_i + b) \mod m.$$

Dabei ist $(\dots) \mod m$ so zu verstehen, dass $x_{i+1} \in \{0, 1, \dots, m-1\}$ der Rest ist, der beim Teilen der rechten Seite durch m entsteht. Die x_i liegen alle in der Menge $\{0, 1, \dots, m-1\}$ und sind periodisch mit einer Periode $\leq m$.

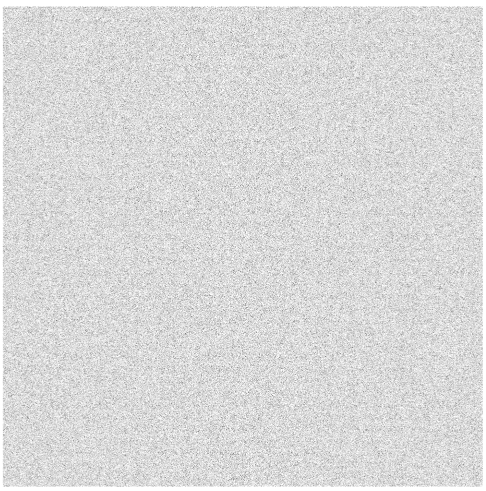
Als einfaches Beispiel nehmen wir $a = 4$, $b = 1$, $m = 11$ und erhalten

$$x_0 = 1,\, x_1 = 5,\, x_2 = 10,\, x_3 = 8,\, x_4 = 0,\, x_5 = 1.$$

Offenbar müssen die Zahlen m, a, b gut gewählt werden, wenn man Periodenlänge m haben will.



Raster 1000 x 1000 Linearer Kongruenzgenerator
a=12453, b=8889, m=247897



Raster 1000 x 1000 IFORT-Compiler
Weiss: kein Treffer Schwarz: Treffer>4

In der Praxis werden die Zahlen m, a, b so bestimmt, dass die Periode maximal, also m ist, und die erzeugten x_i verschiedene Tests auf Zufälligkeit erfolgreich bestehen. Man kann die Folge in

Binärdarstellung hintereinander schreiben. Es soll dann eine zufällige $0, 1$-Folge entstehen, in denen alle Bitmuster in etwa gleich häufig vorkommen.

Man teilt die Folgenglieder x_i des linearen Kongruenzgenerators durch m und erhält (hoffentlich) gleichverteilte Zufallszahlen auf dem Intervall $[0, 1)$. Da die Beschreibung der Folge kurz ist, wird es immer Anwendungen geben, bei denen die „Zufälligkeit" dieser Zahlen nicht ausreicht.

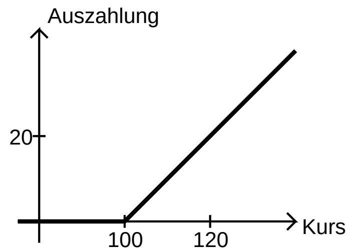
Auch wenn m, a, b im linearen Kongruenz-Generator gut gewählt werden, sind *Zufallsvektoren* problematisch. In diesem Fall bildet man – wie bei der Berechnung von Flächeninhalten – aus den Zufallszahlen die Folge

$$(x_0, x_1),\ (x_2, x_3),\ (x_4, x_5), \ldots .$$

Das Bild links zeigt die von einem linearen Zufallsgenerator erzeugten Vektoren. Obwohl die zugehörigen Zufallszahlen tatsächlich wie zufällig erscheinen, liegen im Bild viele Vektoren auf einer Geraden und sind daher für die Flächenberechnung ungeeignet. Das Bild rechts zeigt die Zufallsvektoren des fortran90-Compilers ifort. Beide Bilder wurden erzeugt, indem ein 1000×1000-Raster in das Einheitsquadrat gelegt und dann gezählt wird, wie oft ein Zufallsvektor in eines der Teilquadrate fällt. Je öfter dies geschieht, desto dunkler wird das Teilquadrat eingefärbt.

Fairer Preis von Finanz-Derivaten Als Beispiel für ein Finanz-Derivat erläutern wir die Option auf eine Aktie. Eine Aktie kostet zum 1.1. eines Jahres 100 Euro. Man kann eine Kaufoption auf diese Aktie erwerben, die einem das Recht gibt, diese Aktie zum 31.12. des gleichen Jahres zum Preis von 100 Euro zu kaufen. Wir nehmen an, dieses Recht kostet 10 Euro. Zum 31.12. kauft man die Aktie zu 100 Euro, sofern der Kurs über 100 Euro liegt, und verkauft sie gleich wieder. Es ergeben sich damit folgende Möglichkeiten zum 31.12.:

Kurs	Gewinn
120	10
100	-10
80	-10



Auszahlungsfunktion einer Kaufoption (ohne Kaufpreis)

Das Bild rechts zeigt die Auszahlungsfunktion dieser Kaufoption. Das Tröstliche daran ist, dass man nicht mehr verlieren kann als seinen Einsatz im Gegensatz zum Aktionär, der mit dieser Aktie im ungünstigsten Fall 100 Euro in den Sand setzt. Daher sieht die Mathematik Optionen durchaus positiv, sofern nur ein kleiner Teil des Vermögens darin angelegt wird.

Auch wenn es mathematisch keinen Unterschied macht, unterscheidet man zwischen Optionen und Optionsscheinen. Erstere werden an Terminbörsen gehandelt und der Preis wird durch Angebot und Nachfrage marktwirtschaftlich ermittelt. Optionsscheine werden dagegen von Handelshäusern begeben, die an der Börse Kauf- und Verkaufskurse stellen. In diesem Fall wird zur Preisfeststellung ein mathematisches Modell benötigt, da ein Markt für die Optionsscheine de facto nicht existiert.

Was ist der faire Wert einer Option? Wir nehmen an, dass der Kurs der Aktie sich im Jahresverlauf zufällig entwickelt mit einer durchschnittlichen Rendite, die der Marktrendite für Anleihen entspricht. Der durchschnittliche Gewinn, der mit einer Option bei diesen zufälligen Kursverläufen erzielt wird, ist gleichzeitig der faire Preis der Option. Gerade für erfahrene Börsianer ist dieser Ansatz erstaunlich, weil die Aktie sich nur besser als der planlose Zufall entwickeln muss, um mit der Option Gewinn zu erzielen.

Wir müssen uns als erstes ein stochastisches Modell des Kursverlaufs verschaffen. Sei I die Laufzeit der Option in Börsentagen und s die maximale relative Kursänderung an einem Börsentag, die aus der Vergangenheit des Kurses bestimmt wird. s beschreibt, wie flatterhaft oder volatil die Aktie in der Vergangenheit gewesen ist. Die Volatilität wird zum einen von der allgemeinen

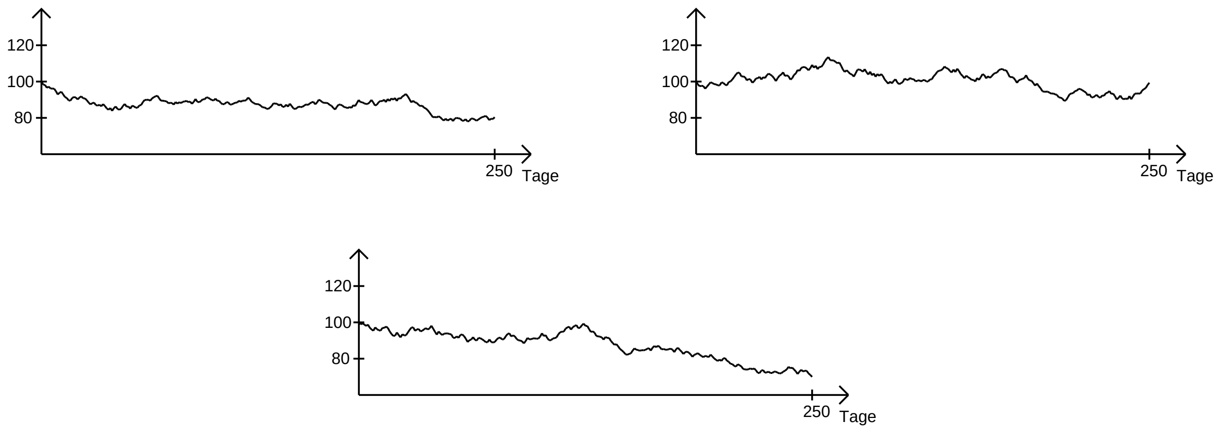
Börsenverfassung bestimmt, zum anderen auch von der Aktiengesellschaft selbst. So überstehen finanzstarke Gesellschaften mit einem wenig anfälligen Geschäftsmodell Krisen besser als Firmen mit einem stark zyklischen Geschäft. Ein Beispiel für Letzteres ist die Stahlindustrie, die bei niedrigem Stahlpreis mit einem Bein im Konkurs steht und bei hohen Preisen finanziell aus dem Vollen schöpft. Die im folgenden Modell unterstellte Hypothese, dass die aus der Vergangenheit bestimmte Volatilität auch für die Laufzeit gültig bleibt, ist also nicht aus der Luft gegriffen. Es ist überdies klar, dass die Volatilität ein wichtiger preisbestimmender Faktor für eine Option darstellt. Denn wenn der Kurs im obigen Beispiel bei 100 Euro verharrt, was der Volatilität $s = 0$ entspricht, wird man seinen Einsatz verlieren.

a sei der Tageszinssatz des Marktes (z.B. $a^{250} = 1.01$ bei etwa 250 Börsentagen) und x_0 der Kurs der Aktie am Tag 0. Mit gleichverteilten Zufallszahlen y_i im Intervall $(-s, s)$ simulieren wir einen zufälligen Aktienkurs durch

$$x_{i+1} = ax_i + y_i x_i, \quad 0 \le i \le I - 1.$$

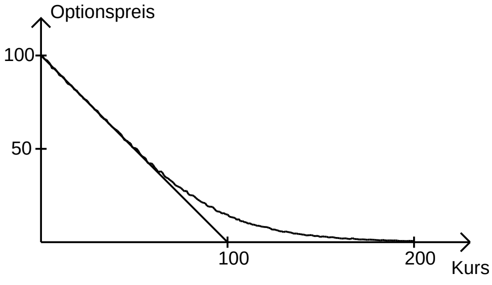
Wenn wir hier $y_i = 0$ setzen, so erhalten wir mit x_i den Tageswert einer Anleihe, die sich in unserem Beispiel mit 1% verzinst. Wenn wir also fortwährend in diesem Modell Aktien kaufen, so erhielten wir im Schnitt am Jahresende 1% Gewinn. Aus x_I kann der Gewinn durch die Option bestimmt werden.

Und nun die Frage an die Leserinnen und Leser. Sind die folgenden Charts „echt“ oder vom Zufallsgenerator erzeugt?



Echte Kurse oder aus dem Zufallsgenerator ?

Wir machen einige 1000 Durchläufe mit den beschriebenen zufälligen Kursverläufen, bestimmen jedesmal x_I und den daraus resultierenden Gewinn und erhalten mit dem Durchschnitt dieser Gewinne den Erwartungswert und damit den fairen Preis für die Option.



Optionspreis einer Verkaufsoption

Im nebenstehenden Bild sehen wir den durch die Simulation bestimmten fairen Preis einer Verkaufsoption bei unterschiedlichen aktuellen Kursen. Die stückweise lineare Funktion ist die Auszahlungsfunktion, die in diesem Zusammenhang als *inneren Wert* der Verkaufsoption bezeichnet wird. Was über dem inneren Wert liegt, wird *Aufpreis* genannt. Der Aufpreis ist offenbar am größten, wenn der aktuelle Kurs der Aktie genau auf dem *Basispreis* von 100 Euro liegt.

Mit der hier vorgestellten Methode lassen sich alle Arten von Derivaten und derivathaltigen Geschäften (auch Genußscheine, Wandelanleihen, Futures, Zertifikate usw.) auf alle Arten von Basisobjekten (Aktienindizes, Währungen, landwirtschaftliche Produkte usw.) untersuchen. Dennoch gilt bei Optionen: Den Optionspreis legt der Markt fest, nicht die Theorie.

Der Verkäufer der Option trägt das Risiko. Mit einer Erweiterung der Theorie lässt sich ein Gegengeschäft des Verkäufers angeben, das ihn im Rahmen der Theorie von jedem Risiko freistellt.

In der Praxis verwendet man normalverteilte an Stelle von gleichverteilten Zufallszahlen und einen kontinuierlichen an Stelle eines diskreten stochastischen Prozesses. In einfachen Fällen kann der Optionspreis durch eine Formel angegeben werden, was man als *Black-Scholes-Theorie* bezeichnet.