

# Glückssträhnen ... ?!

## *Teilnehmer:*

|                      |   |
|----------------------|---|
| Aptin Haerian        | Herder-Oberschule, Berlin                         |
| Max Irmscher         | Johannes-Kepler-Gymnasium, Chemnitz               |
| Markus Johl          | John-Lennon-Oberschule, Berlin                    |
| Felix Montenegro     | Friedrich-Schiller-Gymnasium, Königs Wusterhausen |
| Hoang Lam Nguyen     | Andreas-Oberschule, Berlin                        |
| Anne Christin Rettig | Heinrich-Hertz-Oberschule, Berlin                 |

## *Gruppenleiter:*

|               |   |
|---------------|---|
| Elke Warmuth  | Humboldt-Universität zu Berlin<br>Mitglied im DFG-Forschungszentrum MATHEON<br>„Mathematik für Schlüsseltechnologien“ |
| André Henning | Humboldt-Universität zu Berlin<br>Mitglied im Humboldt-ProMINT-Kolleg   |

Für viele Menschen scheinen bei Glücksspielen sogenannte „Glückssträhnen“, also eine Führung über einen langen Zeitraum, eher unwahrscheinlich. Ein bekanntes Beispiel aus dem Alltag ist das Werfen einer fairen Münze. Fällt Wappen, erhält der eine Spieler einen Punkt, fällt Zahl, bekommt der andere Spieler einen Punkt. Hierbei kann man sich verschiedene Fragen stellen:

- (1) Wie lange wird ein Spieler bei einer großen Anzahl von Spielen führen?
- (2) Wann wird es wieder unentschieden stehen?
- (3) Mit welcher Wahrscheinlichkeit wird das passieren?
- (4) Wie wahrscheinlich sind Glückssträhnen beziehungsweise Pechsträhnen?

Solche Fragen gehören in das Gebiet der Irrfahrten. In diesem Zusammenhang haben wir uns mit sogenannten Markov-Ketten, Absorptionswahrscheinlichkeiten und Rekurrenz beschäftigt.

Mit Hilfe von Java-Simulationen haben wir die Wahrscheinlichkeiten solcher Glückssträhnen experimentell ermittelt.

# 1 Münzbeispiel – Hart, aber fair?

Zur Einführung in unser Thema haben wir zunächst ein Experiment durchgeführt. Dabei wird eine Münze mehrfach hintereinander geworfen. Wir notieren für das Ereignis „es fällt Wappen“ ein  $W$  und für das Ereignis „es fällt Zahl“ ein  $Z$ . Es gewinnt der eine Spieler, wenn  $WWW$  eintritt, der andere Spieler gewinnt, wenn  $WZW$  fällt. Eine mögliche dabei entstandene Kette könnte lauten:

$ZWWZZZWZZWWZW$ .

Das Interessante an diesem Versuch ist nun, dass man annehmen könnte, dass es sich dabei um eine Bernoulli-Kette handelt. Zwar ist die Folge der  $W$  und  $Z$  sehr wohl eine Bernoulli-Kette, aber uns interessiert nicht nur ein einzelner Wurf, sondern uns interessieren Muster in aufeinanderfolgenden Würfeln. Die Wahrscheinlichkeit eines Musters hängt von den vorherigen Würfeln ab. Da wir auf die Muster  $WWW$  oder  $WZW$  warten, kann man in unserer Kette stets diejenigen Glieder wegstreichen, die zu diesen Mustern nichts mehr beitragen. Die senkrechten Striche im folgenden Beispiel veranschaulichen dies:

$Z|W|WZ|Z|Z|WZ|Z|W|WZW$ .

Mit der Zeit läuft so ein stochastischer Prozess ab, der folgende Zustände besitzt:  $W, Z, WW, WZ, WWW, WZW$ . Die möglichen Übergänge zwischen diesen Zuständen und deren Wahrscheinlichkeiten lassen sich mit einem Übergangsgraphen darstellen.

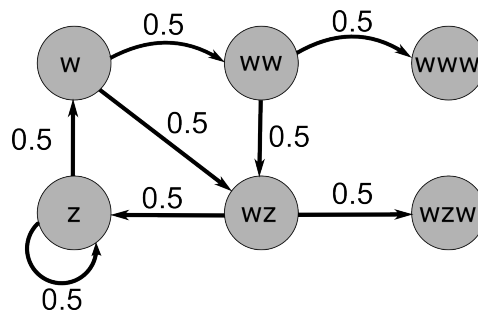


Abbildung 1: Übergangsgraph

Bei unserem Münzwurf-Experiment handelt es sich um ein erstes Beispiel für eine homogene Markov-Kette. Um unser Beispiel näher untersuchen zu können, beschäftigen wir uns nun mit Eigenschaften dieser speziellen stochastischen Prozesse.

## 2 Homogene Markov-Ketten

**Definition 2.1.** Ein stochastischer Prozess  $(X_n)$  mit Zustandsmenge  $I$  heißt homogene Markov-Kette, wenn für alle  $n \in \mathbb{N} = \{0, 1, \dots\}$  und alle  $i_0, i_1, \dots, i_n \in I$  gilt:

$$\begin{aligned} P(X_n = i_n | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}) &= P(X_n = i_n | X_{n-1} = i_{n-1}) \\ &=: p_{i_{n-1}i_n}. \end{aligned}$$

In Worten bedeutet diese Gleichung, dass bei einer Markov-Kette die Zukunft von der Vergangenheit nur über die Gegenwart abhängt.

Eine Markov-Kette ist durch die Übergangswahrscheinlichkeiten  $p_{i_{n-1}i_n}$  und die Anfangsverteilung  $P(X_0 = i)$  eindeutig bestimmt. Die Matrix der Übergangswahrscheinlichkeiten  $p_{ij}$  ( $i, j \in I$ ) wird Übergangsmatrix genannt. Die Übergangswahrscheinlichkeiten für unser Münzwurf-Beispiel kann man Abbildung 1 entnehmen.

### 2.1 Absorptionswahrscheinlichkeiten

Markov-Ketten können in einem Zustand enden. Dies nennt man Absorption. In unserem Münzbeispiel enden die Ketten im Zustand  $WWW$  oder  $WZW$ . Mit Hilfe der Eigenschaften einer Markov-Kette lässt sich die Absorptionswahrscheinlichkeit in  $WWW$  mit Hilfe eines Gleichungssystems berechnen:

$$\begin{aligned} \alpha_W &= \alpha_{WW} \cdot \frac{1}{2} + \alpha_{WZ} \cdot \frac{1}{2} \\ \alpha_{WW} &= \alpha_{WWW} \cdot \frac{1}{2} + \alpha_{WZ} \cdot \frac{1}{2} \\ \alpha_Z &= \alpha_Z \cdot \frac{1}{2} + \alpha_W \cdot \frac{1}{2} \\ \alpha_{WZ} &= \alpha_{WZW} \cdot \frac{1}{2} + \alpha_Z \cdot \frac{1}{2} \end{aligned}$$

$\alpha_W$  ist die Wahrscheinlichkeit in  $WWW$  absorbiert zu werden, wenn man in  $W$  startet. Daher ist  $\alpha_{WWW} = 1$  und  $\alpha_{WZW} = 0$ . Das resultierende Gleichungssystem hat die Lösung  $\alpha_W = \alpha_Z = \frac{2}{5}$ . Damit ist die Absorptionswahrscheinlichkeit unabhängig vom ersten Wurf gleich  $0,4$ .

### 2.2 Rekurrenz

Im Münzspiel muss der Anfangszustand nicht unbedingt wieder erreicht werden. Beispielsweise kann es passieren, dass die Kette vor dem erneuten Erreichen des Anfangszustandes absorbiert wird. Man sagt, der Anfangszustand ist nicht rekurrent.

**Definition 2.2.** Ein Zustand  $i \in I$  einer Markov-Kette heißt rekurrent, wenn er mit Wahrscheinlichkeit 1 jemals wieder erreicht wird.

Für die Untersuchung der Rekurrenz sind folgende Wahrscheinlichkeiten von Bedeutung:

- $p_n = P(X_n = i | X_0 = i)$  ist die Wahrscheinlichkeit dafür, dass zum Zeitpunkt  $n$  der Anfangszustand besucht wird.
- $p^* = \sum_{n=1}^{\infty} p_n$  ist der Erwartungswert der Anzahl der Besuche im Anfangszustand.
- $f_n = P(X_1 \neq i, X_2 \neq i, \dots, X_{n-1} \neq i, X_n = i | X_0 = i)$  ist die Wahrscheinlichkeit dafür, dass zum Zeitpunkt  $n$  der Anfangszustand zum ersten Mal wieder besucht wird.
- $f^* = \sum_{n=1}^{\infty} f_n$  ist die Wahrscheinlichkeit dafür, dass der Anfangszustand jemals wieder erreicht wird.
- $q_n = P(i \text{ wird genau } n\text{-mal besucht} | X_0 = i)$ .

**Satz 2.1.** Für die Anzahl  $B_i$  der Besuche in  $i$  gilt:

$$P(B_i \geq m) = (f^*)^m. \quad (2.1)$$

*Beweis.* Wir beweisen 2.1 durch vollständige Induktion.

Per Definition gilt  $P(B_i \geq 1) = f^*$ .

Angenommen, für eine natürliche Zahl  $m$  gilt 2.1. Es folgt:

$$\begin{aligned} P(B_i \geq m + 1) &= P(B_i \geq m + 1, B_i \geq m) \\ &= P(B_i \geq m + 1 | B_i \geq m) \cdot P(B_i \geq m) \\ &= f^* \cdot (f^*)^m \\ &= (f^*)^{m+1}. \end{aligned}$$

□

Der folgende Satz stellt eine Verbindung zwischen  $p^*$  und  $f^*$  her, die es ermöglicht die Rekurrenz mit Hilfe von  $p^*$  zu untersuchen, was im Allgemeinen einfacher ist, als  $f^*$  direkt zu bestimmen.

**Satz 2.2.** Es gilt  $f^* = 1$  genau dann, wenn  $p^* = +\infty$ .

*Beweis.* Wir finden zunächst eine nützliche andere Darstellung für den Erwartungswert  $E(B_i)$ , die für beliebige Zufallsgrößen mit Werten in  $\mathbb{N}$  gilt:

$$p^* = E(B_i) = \sum_{k=1}^{\infty} k \cdot P(B_i = k) = \sum_{k=1}^{\infty} k \cdot q_k = \sum_{k=1}^{\infty} \left( \sum_{j=0}^{\infty} q_{k+j} \right) = \sum_{k=1}^{\infty} P(B_i \geq k).$$

Gemäß 2.1 erhalten wir:

$$p^* = \sum_{k=1}^{\infty} (f^*)^k.$$

Angenommen, es gilt  $f^* = 1$ . Dann folgt  $p^* = \sum_{k=1}^{\infty} 1^k = \infty$ .

Um die Umkehrung zu beweisen, zeigen wir: Aus  $f^* < 1$  folgt  $p^* < +\infty$ . Da  $0 < f^* < 1$ , konvergiert die geometrische Reihe und wir erhalten:

$$p^* = \frac{f^*}{1 - f^*} < +\infty.$$

□

### 3 Irrfahrten

Irrfahrten sind spezielle Markov-Ketten. Der Übergangsgraph einer eindimensionalen Irrfahrt ist durch folgendes Bild gegeben. Die Wahrscheinlichkeit, von einem Zustand  $i$  in den Zustand  $i + 1$  zu gelangen, ist  $p$ .

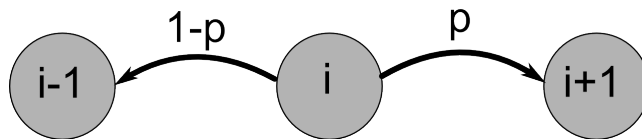


Abbildung 2: Irrfahrtgraph

$Y_k$  ist die Änderung in  $y$ -Richtung im  $k$ -ten Schritt.

$$Y_k = \begin{cases} 1 & \text{mit Wahrscheinlichkeit } p \\ -1 & \text{mit Wahrscheinlichkeit } 1 - p \end{cases}, k \in \mathbb{N}.$$

Diese Schritte sind alle voneinander unabhängig. Summiert man sie auf, erhält man den Zustand zu einem bestimmten Zeitpunkt:

$$S_0 = 0, S_n = Y_1 + Y_2 + \dots + Y_n.$$

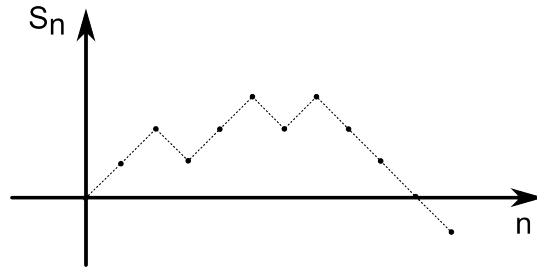


Abbildung 3: Irrfahrtpfad

Den zeitlichen Ablauf einer Irrfahrt, die in 0 startet, kann man wie in Abbildung 3 veranschaulichen.

Die Irrfahrt  $(S_n)$  in  $\mathbb{R}$  ist eine Markov-Kette mit der Übergangsmatrix:

$$p_{ij} = \begin{cases} p & j = i + 1 \\ 1 - p & j = i - 1 \\ 0 & \text{sonst} \end{cases} .$$

Irrfahrten kann man in beliebigen Dimensionen von  $\mathbb{R}$  betrachten. Wir beschränken uns jedoch auf den  $\mathbb{R}^1$ .

### 3.1 Rückkehrwahrscheinlichkeit

Bei einer Irrfahrt ist es interessant, die Wahrscheinlichkeit zu untersuchen, mit der der Pfad zum Ursprungszustand zurückkehrt. Hierzu werden wir  $p^*$  berechnen, um eine Aussage über  $f^*$  treffen zu können. Zuerst berechnen wir die Wahrscheinlichkeit  $p_{2n}$ , da der Pfad nur nach einer geraden Anzahl von Schritten an den Ursprung zurückkehren kann:

$$p_{2n} = P(S_{2n} = 0).$$

Dazu müssen gleich viele Schritte in positive, wie in negative  $y$ -Richtung zurückgelegt werden. Für die  $n$  Schritte in positive  $y$ -Richtung besteht jeweils die Wahrscheinlichkeit  $p$ . Um diese  $n$  Schritte aus allen  $2n$  Schritten auszuwählen, benutzen wir den Binomialkoeffizienten:

$$\begin{aligned} p_{2n} &= p^n (1-p)^n \binom{2n}{n} \\ &= (p(1-p))^n \frac{2n!}{n!n!}. \end{aligned}$$

Um  $p^*$  zu berechnen bilden wir zunächst die Reihe der  $p_{2n}$ :

$$p^* = \sum_{n=1}^{\infty} p_{2n} = \sum_{n=1}^{\infty} (p(1-p))^n \frac{2n!}{n!n!}.$$

Nach Einsetzen der Stirlingschen Formel (3.2) ergibt sich:

$$\sum_{n=1}^{\infty} (p(1-p))^n \frac{2^{2n}}{\sqrt{\pi n}} = \frac{1}{\sqrt{\pi}} \sum_{n=1}^{\infty} (4p(1-p))^n \frac{1}{\sqrt{n}}.$$

Für den Fall  $p \neq \frac{1}{2}$  gilt  $0 < 4p(1-p) < 1$ . Wir setzen  $q := 4p(1-p)$ . Dann erhalten wir für die Reihe folgende Abschätzung:

$$\frac{1}{\sqrt{\pi}} \sum_{n=1}^{\infty} q^n \frac{1}{\sqrt{n}} \leq \frac{1}{\sqrt{\pi}} \sum_{n=1}^{\infty} q^n.$$

Diese geometrische Reihe konvergiert, da  $0 < q < 1$ . Wie oben bewiesen, gilt:

$$p^* \neq \infty \Leftrightarrow f^* < 1.$$

Es ist somit nicht gesichert, dass der Pfad bei einer asymmetrischen Irrfahrt jemals zu seinem Ursprungszustand zurückkehrt.

Für den Fall  $p = \frac{1}{2}$  gilt  $q = 1$ . Damit erhalten wir für die Reihe folgende Abschätzung:

$$\frac{1}{\sqrt{\pi}} \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \leq \frac{1}{\sqrt{\pi}} \sum_{n=1}^{\infty} \frac{1}{n}.$$

Die harmonische Reihe divergiert. Es gilt  $p^* = \infty \Leftrightarrow f^* = 1$ . Der Pfad wird bei einer symmetrischen Irrfahrt mit Sicherheit in den Ursprungszustand zurückkehren.

## 3.2 Stirlingsche Formel

Die Stirlingsche Formel liefert eine asymptotische Näherung für  $n!$ . Diese Näherung ist äußerst hilfreich, da sich mit  $n!$  nur schwer arbeiten lässt.

Die Stirlingsche Formel lautet:

$$n! \underset{\text{as}}{\sim} \sqrt{2\pi n} n^n e^{-n},$$

wobei  $a_n \underset{\text{as}}{\sim} b_n$  bedeutet, dass  $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$  ist.

Wir haben die Formel in der Gruppe nur für ein festes positives  $c$  bewiesen, ohne zu zeigen, dass es sich dabei um den Faktor  $\sqrt{2\pi}$  handelt. Einen vollständigen Beweis findet man zum Beispiel in [2]. Wir haben diese Formel bei Irrfahrten benutzt und mit Hilfe der Formel Wahrscheinlichkeiten abgeschätzt.

### 3.3 Warten auf das Unentschieden

Wir wollen den Erwartungswert  $E(T_0)$  des Zeitpunkts  $T_0$  berechnen, zu dem die Irrfahrt erstmals wieder die Abszisse berührt. Wir wissen über Erwartungswerte, dass sie sich durch die Summe der Anzahlen multipliziert mit den zugehörigen Wahrscheinlichkeiten berechnen lassen. Somit ergibt sich folgende Formel:

$$\begin{aligned} E(T_0) &= \sum_{n=1}^{\infty} 2n \cdot P(T_0 = 2n) \\ &= \sum_{n=1}^{\infty} 2n \cdot f_{2n} \quad (\text{Definition von } f_{2n}). \end{aligned}$$

Für  $f_{2n}$  haben wir einen Zusammenhang mit  $p_{2n-2}$  bewiesen und das Ergebnis eingesetzt:

$$\begin{aligned} E(T_0) &= \sum_{n=1}^{\infty} 2n \cdot f_{2n} \\ &= \sum_{n=1}^{\infty} 2n \cdot \frac{1}{2n} \cdot p_{2n-2}. \end{aligned}$$

Elementare Umformungen führen schließlich zu:

$$\begin{aligned} E(T_0) &= 1 + \sum_{n=1}^{\infty} p_{2n} \\ &= 1 + p^* \\ &= +\infty. \end{aligned}$$

Obwohl wir mit der Wahrscheinlichkeit 1 zum Anfangszustand zurückkehren, ist der Erwartungswert für die erste Rückkehr  $+\infty$ . In der Sprache eines Spielers bedeutet das, dass man theoretisch sehr lange auf das Unentschieden warten kann.

## 4 Glücksstrahlen

Wir haben mit Hilfe von BlueJ (Java)  $N = 10^6$  Irrfahrten mit einer festen Länge von  $2n = 50$  simuliert. Dabei wurden bei jedem Durchgang Zufallszahlen generiert um Irrfahrten zu erzeugen. Anhand dieser Irrfahrten konnte anschließend die Verteilung der Aufenthaltsdauer des Pfades oberhalb der  $x$ -Achse ermittelt werden. Oberhalb der  $x$ -Achse bedeutet hierbei, dass der Pfad sich die ganze Zeit oberhalb der  $x$ -Achse oder in einem Zustand auf der  $x$ -Achse befindet. Die Simulation ergab folgende Häufigkeitsverteilung der Zeiten  $t_k$  des Aufenthaltes oberhalb der  $x$ -Achse. In der Tabelle gibt  $h_k$  die relative Häufigkeit des Wertes  $t_k$  aus  $N$  Simulationen an.

|                   |       |       |       |       |       |       |       |       |       |       |       |       |       |
|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $\frac{t_k}{h_k}$ | 0     | 2     | 4     | 6     | 8     | 10    | 12    | 14    | 16    | 18    | 20    | 22    | 24    |
|                   | 0,122 | 0,057 | 0,044 | 0,037 | 0,033 | 0,030 | 0,029 | 0,028 | 0,027 | 0,026 | 0,026 | 0,025 | 0,025 |
| $\frac{t_k}{h_k}$ | 26    | 28    | 30    | 32    | 34    | 36    | 38    | 40    | 42    | 44    | 46    | 48    | 50    |
|                   | 0,025 | 0,025 | 0,025 | 0,026 | 0,027 | 0,028 | 0,029 | 0,031 | 0,034 | 0,037 | 0,044 | 0,058 | 0,112 |



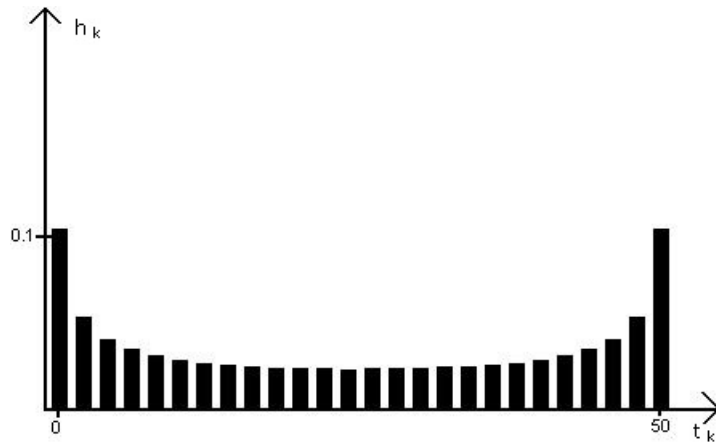


Abbildung 4: Simulationsergebnis  $2n = 50$

Intuitiv könnte man denken, dass eine sogenannte „Glückssträhne“, also eine sehr hohe Aufenthaltsdauer oberhalb der  $x$ -Achse, eher unwahrscheinlich für den Spieler ist (beziehungsweise eine „Pechsträhne“ für den anderen Spieler). Anhand des Diagramms kann man jedoch erkennen, dass Glückssträhnen für beide Spieler besonders wahrscheinlich sind. Am unwahrscheinlichsten ist ein ausgeglichenes Spiel, bei dem der Pfad also genauso lange oberhalb der  $x$ -Achse verläuft wie unterhalb.

## Literatur

- [1] W. Feller, *An Introduction to Probability Theory and Its Applications*, Volume I, John Wiley & Sons, Inc., New York, 1957.
- [2] U. Krengel, *Einführung in die Wahrscheinlichkeitstheorie und Statistik*, Vieweg+Teubner, 2000, S. 89.

