Schuljahr 2024/25
Klassenstufen 9/10

Stochastik - Das Chow-Robbins Spiel

Dies ist eine Zusammenfassung der zwei Einheiten zu stochastischen Spielen. Konkret ging es um Das Chow-Robbins Spiel. Die perfekte Strategie für das Spiel zu finden, ist ein immernoch teilweise ungelöstes Problem. Das Spiel selbst ist jedoch einfach zu verstehen, und wir nutzen es hier, um ein paar Einblicke in mathematische Forschung zu geben.

1 Einführung

Spielregeln: Wirf eine Münze so oft du willst. Wenn du aufhörst, wird dir als Gewinn der Anteil an *Kopf*-Würfen an den Gesamtwürfen ausgezahlt. Z. B. bei 1 Mal *Kopf* aus 2 Würfen 50 Cent, bei 3 Mal *Kopf* aus 5 Würfen 60 Cent.

Aufgabe 1 (a) Im ersten Wurf ist Kopf gefallen. Werft ihr weiter, oder beendet ihr das Spiel? Welches ist die optimale Strategie, also was maximiert euren durchschnittlichen Gewinn?.

(b) Wie sieht es aus, wenn im ersten Wurf Zahl fällt?

(c) Wie ist es bei *ZK*, also erst Zahl, dann Kopf?

(d) Wie ist es bei *ZKK*. Hier müsst ihr Raten.

Schätzfrage Was wäre ein fairer Einsatz für diese Spiel? Ratet und begründet eure Wahl, wenn ihr wollt.

Ein fairer Einsatz wäre der durchschnittliche Gewinn bei optimalem Spiel. Diesen Wert nennen wir W. Wir können feststellen, dass $W \geqslant 0.75$, denn mit der einfachen Strategie "wenn im Ersten Wurf *Kopf* fällt stoppe, sonst warte bis der Gewinn 0.5 ist" gewinnen wir durchschnittlich $\frac{1}{2}(1 + 0.5) = 0.75$. Wir können aber bessere Strategien finden. Tatsächlich ist $W \approx 0.793$.

2 Modellierung

Um das Spiel besser mathematisch untersuchen zu können reformulieren wir es etwas. Im Folgenden wird es etwas technischer und ohne stochastische Vorbildung ggf. etwas schwieriger nachzuvollziehen.
Die Idee ist, statt nur *Kopf*-Würfe zu zählen, verwenden wir einen *Random Walk*, d.h. wir werfen in jeder Runde eine Münze, wenn *Kopf* fällt laufen wir einen Schritt nach oben, bei *Zahl* einen nach unten. Formal bedeutet das:

- Seien X_i unabhängige Zufallsvariablen mit $P(X_i = -1) = P(X_i = 1) = \frac{1}{2}$
- $S_n := \sum_{k=i}^{n} X_i$ bezeichnen wir als *Random Walk*. Die Auszahlung des transformierten Spiels ist nun $\frac{S_n}{n}$ (ein wert zwischen -1 und 1).
- Spielstrategien bezeichnen wir mit *Stoppzeiten* τ, die optimale Strategie mit τ^*.
- Der durchschnittliche Gewinn mit Strategie τ nennen wir W_τ

Erste wichtige Feststellungen sind:

- Das Problem ist lösbar, d.h. es existiert eine optimale Strategie.
- Die Strategie ist nur abhängig vom *Zustand*, nicht vom *Weg*.Das heißt, die optimale Strategie ist eine Linie. Es ist optimal zu stoppen, wenn wir diese überspringen.

Eine typische Lösungsstrategie für solche unendlichen Probleme ist, sie künstlich endlich zu machen. Dafür wählen wir einen endlichen Zeithorizont $T \in \mathbb{N}$. In T endet das Spiel. Wir ändern das Spiel noch so, dass wir immer mindestens eine Auszahlung von 0 bekommen, da wir nach dem *Gesetz der großen Zahlen* immer warten könnten, bis wir 0 bekommen. Von T können wir nun per *Rückwärtsinduktion* die Strategie für die vorherigen Zeitpunkte berechnen.

Aufgabe 2 Bestimmt die perfekte Strategie für das Spiel mit endlichem Zeithorizont $T = 5$. Was ist der durchschnittliche Gewinn?

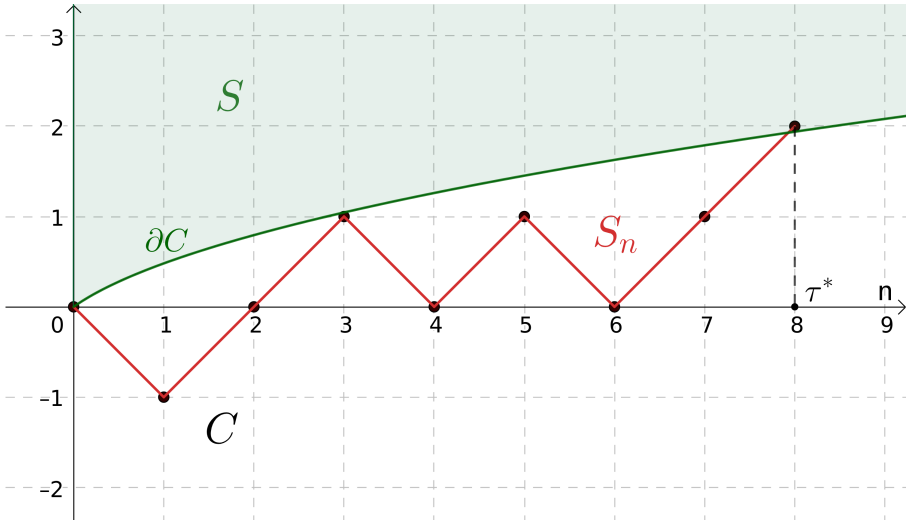


Figure 1: Das Setting des Chow-Robbins Spiels. Die Strategie ist zu stoppen, wenn der Random Walk das erste mal über die Grenze ∂C in den grünen Stoppbereich springt.

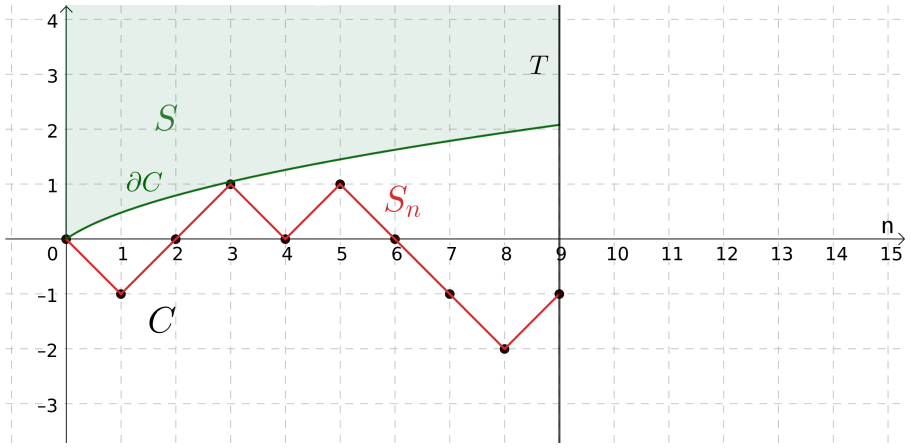


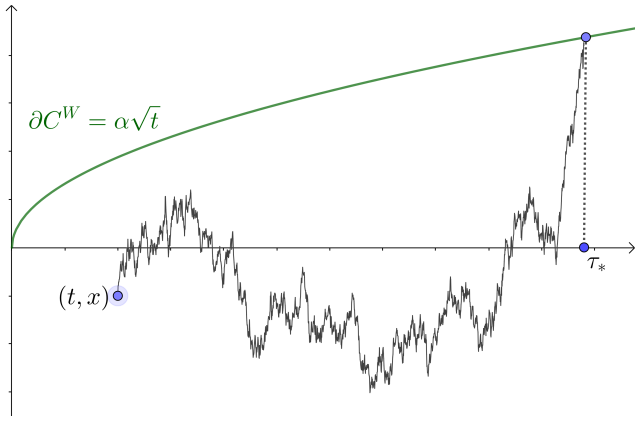
Figure 2: Das Spiel mit endlichem Zeithorizont T

3 Forschung

Es folgt eine sehr knappe Zusammenfassung von Forschungsideen. Für detaillierte Ausführungen siehe [CF22].
Durch Einführung des endlichen Zeithorizonts T bekommen wir eine untere Schranke für die Stoppgrenze. Durch $T \to \infty$ wird diese immer besser. Um eine obere Schranke zu erhalten betrachten wir das Problem mit eine standard Bronwschen Bewegung als Zufallsprozess, statt dem Random Walk. Für diesen ist die Lösung des Problems bekannt:

Satz 1 Die optimale Strategie für das Chow-Robbins Spiel mit Brownscher Bewegung W_t als Prozess ist:

Stoppe wenn das erste mal $W_t > \alpha\sqrt{t}$. *Dabei ist* $\alpha \approx 0{,}83992$.



Eine obere Schranke für die Stoppgrenze (und den Wert des Spiels W) erhalten nur durch folgende Konstruktion. Spiele das ursprüngliche Spiel bis maximal T, dannach ändern sich die Regeln und der Random Walk wird durch eine standard Brownsche Bewegung ersetzt (diese behandeln wir an einem anderen Termin).

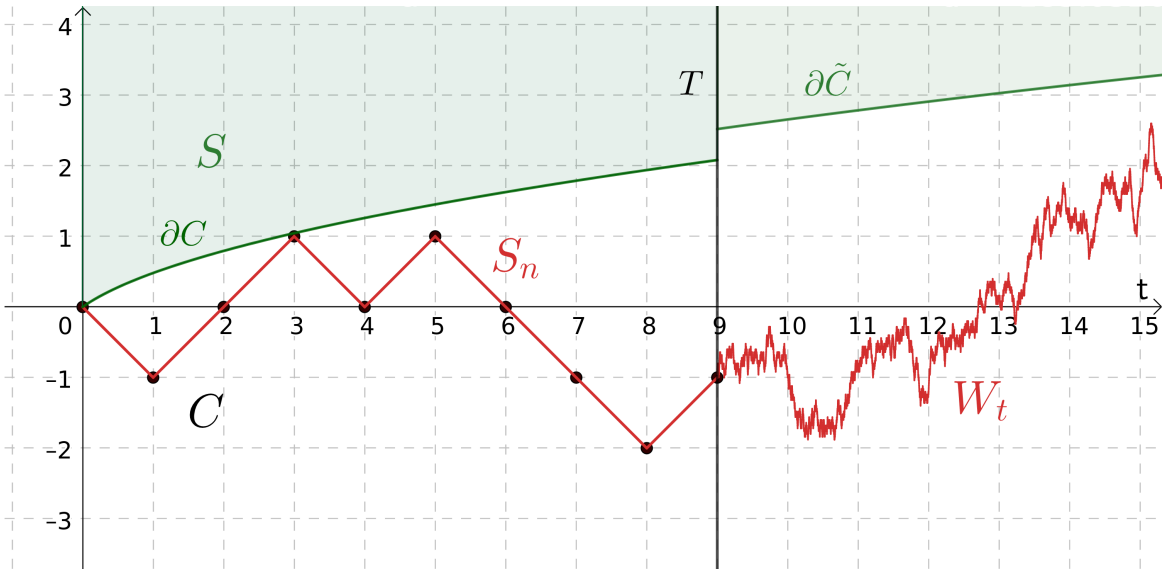


Figure 3: Wechsel zur Brownschen Bewegung W_t ab T

Mit dieser Technik (und einigen anderen) erhalten wir schließlich:

Satz 2 (Christensen, Fischer 2022) *Für ganzzahlige* $n \leqslant 489.241$ *ist die Stoppgrenze* b *gegeben durch*

$$b(n) = \left\lceil \alpha\sqrt{n} - \frac{1}{2} + \frac{1}{7{,}9 + 4{,}54\sqrt[4]{n}} \right\rceil \tag{1}$$

mit den 8 Ausnahmen:

n	b(n)	n	b(n)	n	b(n)	n	b(n)
3195	*48*	*14312*	*101*	*25257*	*134*	*51434*	*191*
12923	*96*	*24880*	*133*	*44653*	*178*	*116342*	*287*

Der durchschnittliche Gewinn ist

$$0{,}5859070128172 \leqslant W_{\tau^*} \leqslant 0{,}5859070128182.$$

4 Lösungen

Lösung 1 (a) Es ist optimal zu stoppen, denn der aktuelle Gewinn ist mit 1 bereits maximal. Es kann nur schlechter werden.

(b) Es ist optimal weiterzuspielen, denn der aktuelle Gewinn ist mit 0 maximal. Es kann nur besser werden.

(c) Es ist optimal weiterzuspielen. Eine Auszahlung von 0.5 können wir langfristig immer bekommen, das garantiert uns das Gesetz der großen Zahlen.[1]

(d) Es ist (sehr knapp) besser weiterzuspielen. Dies kann (nur) durch Rückwärtsinduktion von einem endlichen Zeithorizont gezeigt werden.

Lösung 2 Wir gehen immer einen Zeitschritt zurück. In $n = 4 = T - 1$ berechnen wir den erwarteten Gewinn für jeden Spielstand, wenn wir weiterspielen. Bei $S_4 = 4$ wäre das $\frac{1}{2}\left(1 + \frac{3}{5}\right) = \frac{4}{5}$, was kleiner ist als die Auszahlung von 1, wenn wir stoppen, wir stoppen also. Bei $S_4 = 0$ bekommen wir beim Weiterspielen $\frac{1}{2}\left(\frac{1}{5} + 0\right) = \frac{1}{10}$, was größer ist als die Auszahlung von 0, wenn wir stoppen, wir spielen also weiter. Für $n = 3$ berechnen wir entsprechend unseren erwarteten Gewinn, wenn wir in $n = 4$ optimal spielen, usw.
Es ist optimal zu stoppen, wenn
$n = 1$, $S_n \geqslant 1$
$n = 2$, $S_n \geqslant 2$
$n = 3$, $S_n \geqslant 1$
$n = 4$, $S_n \geqslant 0$
Der erwartete Gewinn des Spiels ist 0.55416, das entspricht einem erwarteten Gewinn von 0.777083 im ursprünglichen (unsymmetrischen) Spiel.

References

[CF22] Sören Christensen and Simon Fischer. On the sn/n problem. *Journal of Applied Probability*, 59:1–13, 05 2022.

[1] Für die mit stochastischer Vorbildung: Genaugenommen garantiert das GdgZ nur, dass wir beliebig nahe an die 0.5 herankommen. Dass wir die 0.5 auch erreichen, garantiert uns die *Rekkurenz* des *Random Walks*.