Vortrag zum Thema

Monte Carlo Tennis

gehalten von

Andreas Bender

am

5. Dezember 2009

im Rahmen des Bachelor Seminars

Statistik im Sport

bei Prof. Dr. Friedrich Leisch, Manuel Eugster, Sebastian Kaiser
am Institut für Statistik
Ludwidg–Maximilians–Universität
München

Spielweise

- 1 vs. 1 (Einzel) oder 2 vs. 2 (Doppel)
- Bestimmung des ersten Aufschlägers durch Los
- Match ist unterteilt in Punkte, Spiele und Sätze
- Sieger ist derjenige Spieler, der eine vorher festgelegte Anzhal Sätze gewonnen hat
- Meistens 2 Gewinnsätze bei Frauen, 3 bei Männern

Zählweise

Punkte

- Normalerweise haben Punkte im Tennis spezielle Bezeichnungen

kein Punkt	'0', 'Love'
1. Punkt	'15'
2. Punkte	'30'
3. Punkte	'40'
3:3	'Gleichstand', 'Deuce'
Punkt nach 3:3	'Vorteil', 'Advantage'

- Hier soll von $0, 1, 2, \ldots$ gezählt werden

Zählweise

Spiele

- In einem Spiel schlägt immer nur ein Spieler auf, bis das Spiel gewonnen oder verloren ist
- Ein Spiel ist gewonnen oder verloren
  - beim Spielstand $4 : x$ bzw. $x : 4$ mit $x \leq 2$, oder
  - nach Gleichstand, wenn ein Spieler 2 Punkte Vorsprung hat
- Nach jedem Spiel wechselt das Aufschlagsrecht
- Gewonnene Spiele werden zusammengezählt, bis einer der Spieler den Satz gewinnt

Zählweise
Sätze

- Ein Satz ist gewonnen, wenn es nach Spielen
  - $6 : x$ bzw. $x : 6$, mit $x \leq 4$, oder $7 : 5$ steht
  - kommt es zum Gleichstand($6 : 6$), wird der Sieger im Tiebreak ausgespielt
- Das Tiebreak ist gewonnen oder verloren beim Punktestand
  - $7 : x$ bzw. $x : 7$, mit $x \leq 5$, oder
  - nach Gleichstand ($6 : 6$) mit 2 Punkten Vorsprung
  - Erster Aufschläger im Tiebreak $\widehat{=}$ Rückschläger des vorangegangenen Spiels
- Das Aufschlagsrecht für das erste Spiel eines Satzes alterniert
- Ein Spieler siegt mit der erforderlichen Anzahl Gewinnsätze

Analytische Wahrscheinlichkeiten
Notation

Im Folgenden soll die untenstehende Notation gelten:

- $p_A^R \widehat{=}$ Wahrscheinlichkeit für einen Punktgewinn durch Spieler A bei eigenem Aufschlag ('rally' = engl. für Ballwechsel)
- $p_A^G \widehat{=}$ Wkt. für ein Spielgewinn durch Spieler A, wenn A Aufschläger ('game' = engl. für Spiel)
- $p_A^T \widehat{=}$ Wkt. für den Gewinn des Tiebreak durch Spieler A, wenn A erster Aufschläger
- $p_A^S \widehat{=}$ Wkt. für einen Satzgewinn durch Spieler A, wenn A erster Aufschläger im Satz ('set' = engl. für Satz)
- $p_A^M \widehat{=}$ Wkt. für den Gewinn des Matches durch Spieler A, wenn A erster Aufschläger im Match

Analytische Wahrscheinlichkeiten
Spielgewinnwahrscheinlichkeit

- p_A^G = Summe der Wahrscheinlichkeiten der verschiedenen Spielausgänge, die A als Sieger hervorbringen (z. B. $4 : 0$; $4 : 2$; $9 : 7$; etc.)
- Mit $q_A^R = 1 - p_A^R$ ergibt sich dann:

$$p_A^G = (p_A^R)^4[1+4q_A^R+10(q_A^R)^2]+20(p_A^R q_A^R)^3(p_A^R)^2[1-2p_A^R q_A^R]^{-1}$$

- Die ganzzahligen Faktoren stehen dabei für die verschiedenen Kombinationen, mit denen der jeweilige Spielstand eintreten kann
- Man beachte, dass p_A^G nur von der Punktgewinnwahrscheinlichkeit von Spieler A (p_A^R) abhängt.

Analytische Wahrscheinlichkeiten
Spielgewinnwahrscheinlichkeit (Beispiel 4 : 1)

- Betrachte Punktestand $4 : 1$. Dann macht
  - Spieler A vier Punkte $\longrightarrow (p_A^R)^4$
  - Spieler B einen Punkt $\longrightarrow (q_A^R)^1$
- Dafür gibt es folgende Möglichkeiten:

1. Punkt	2. Punkt	3. Punkt	4. Punkt	5. Punkt
A	A	A	B	A
A	A	B	A	A
A	B	A	A	A
B	A	A	A	A

- B hat also $\binom{4}{1} = 4$ Möglichkeiten einen Punkt zu machen
- Damit ist $\mathbb{P}(4 : 1) = 4 \cdot (p_A^R)^4 (q_A^R)^1$

Analytische Wahrscheinlichkeiten

Satzgewinnwahrscheinlichkeit

- $p_A^S \mathrel{\widehat{=}}$ Summe der Wahrscheinlichkeiten der verschiedenen Satzgewinnmöglichkeiten durch A (z. B. 6 : 0; 6 : 4; 7 : 6; etc.)
- Dann gilt allgemein:

$$p_A^S = \sum_{j=0}^{4} p_A^S(6,j) + p_A^S(7,5) + p_A^S(6,6)p_A^T$$

- Sei $q_A^G = 1 - p_A^G$ und $q_B^G = 1 - p_B^G$
- Dann gilt ähnlich zur Spielgewinnwahrscheinlichkeit:
  - $p_A^S(6:0) = (p_A^G q_B^G)^3$
  - $p_A^S(6:1) = 3(p_A^G)^3 q_A^G (q_B^G)^3 + 3(p_A^G)^4 p_B^G (q_B^G)^2$
  - ...

Analytische Wahrscheinlichkeiten

Matchgewinnwahrscheinlichkeit

- Wahrscheinlichkeit für einen Sieg hängt von der Anzahl Gewinnsätze ab
- Für 2 Gewinnsätze ergibt sich

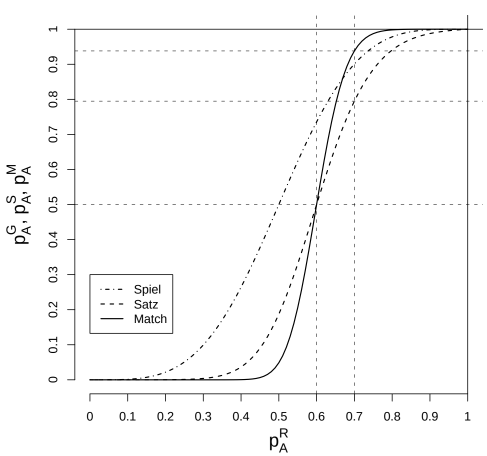
$$p_A^M = (p_A^S)^2 + 2(p_A^S)^2 p_B^S$$

- Für 3 Gewinnsätze folgt

$$p_A^M = (p_A^S)^3 + 3(p_A^S)^3 p_B^S + 6(p_A^S)^3 (p_B^S)^2$$

Analytische Wahrscheinlichkeiten

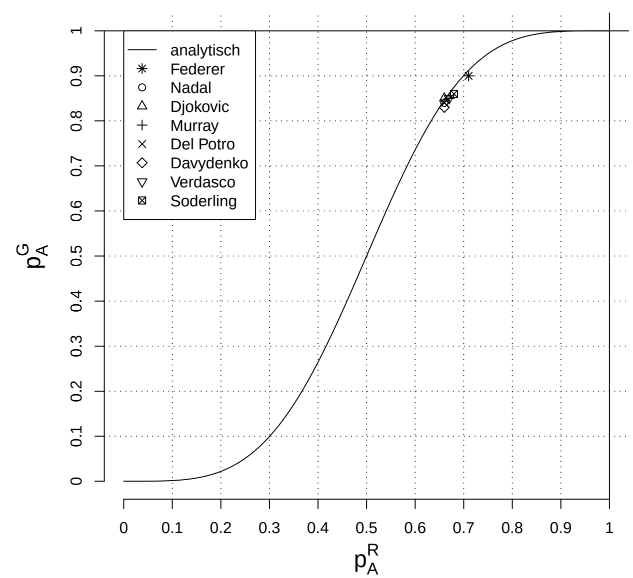
Graphische Darstellung



- Zuhnemend steile Kurvenverläufe in der Reihenfolge Spiel, Satz, Match
- Zählweise begünstigt besseren Spieler
- Wahrscheinlichkeiten p_A^S und p_A^M nicht vom ersten Aufschläger abhängig

Exkurs

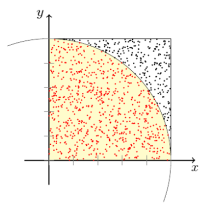
Vergleich mit empirischen Werten



Das Monte–Carlo–Verfahren

Überblick

- Numerisches Verfahren zur Lösung deterministischer und stochastischer Probleme
- Anwendungsgebiete:
  - Berechung von Integralen, Summen, Gleichungssysteme, ...
  - Warteschlangenprobleme, Lagerhaltungsprobleme, atomare Prozesse usw.
- Klassisches Beispiel: Schätzen von π



Monte–Carlo–Verfahren

Vorgehen

1. Möglichst genaue stochastische Modellierung der jeweiligen Problemstellung $Y = f(X_1, ..., X_n)$
2. Berechnen der (ggf. zusammengesetzten) Zufallsvariable y_i aus der durch Zufallszahlen simulierten Stichprobe $y_i = f(x_{i_1}, ..., x_{i_n})$
3. n-faches Wiederholen von Schritt 2. und Auswertung von $y = (y_1, ..., y_n)$

Monte–Carlo–Verfahren

Anwendung auf Tennis

- Ziehen von gleichverteilten Zufallszahlen u_i aus $[0, 1]$

$$X_i = \begin{cases} 1, & u_i \in [0, p], \text{ Punkt Aufschläger} \\ 0, & \text{sonst, Punkt Rückschläger} \end{cases}$$

- Zählen von Punkten, Spielen und Sätzen bis einer der Spieler ein Spiel, Satz bzw. Match gewinnt
- n-faches Simulieren von Spielen, Sätzen und Matches
- Schätzen von p^G, p^S, p^M durch die entsprechenden Stichprobenmittel:

$$\widehat{p^G} = \frac{1}{n}\sum_{i=1}^{n} y_i^G,\ \widehat{p^S} = \frac{1}{n}\sum_{i=1}^{n} y_i^S,\ \widehat{p^M} = \frac{1}{n}\sum_{i=1}^{n} y_i^M$$

Monte–Carlo–Verfahren

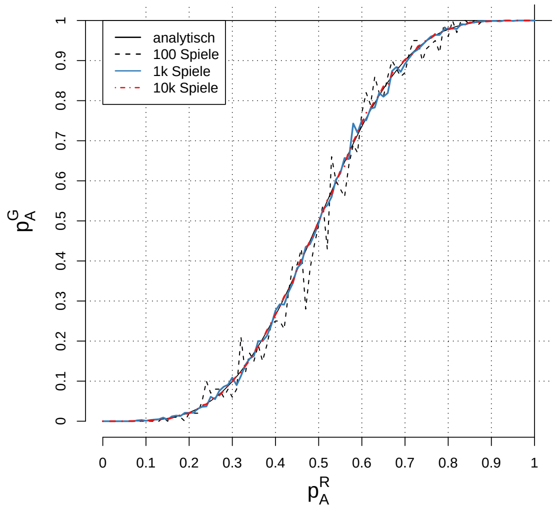
Beispiel: Schätzung von p_A^G

Angenommen wir wollen p_A^G durch Simulation von Spielen schätzen:

- Sei $p_A^R = 0.7$ und die gezogenen Zufallszahlen $(0.4, 0.12, 0.9, 0.7, 0.5)$
- Dann ist $x_1 = (1, 1, 0, 1, 1)$, d. h. die Simulation erreicht einen Punktestand von 4 : 1 und A gewinnt sein Aufschlagsspiel (also $y_{A_1}^G = 1$)
- Seien insgesamt 10 Spiele simuliert mit bspw. $y_A^G = (y_{A_1}^G, ..., y_{A_{10}}^G) = (1, 0, 1, 1, 1, 0, 0, 1, 0, 1)$
- Dann ist der MC-Schätzer $\widehat{p_A^G} = \bar{y}_A^G = \frac{6}{10} = 0.6$
- Analytisch: $p_A^G(0.7) \approx 0.90$

Vergleich Simulation vs. Theorie

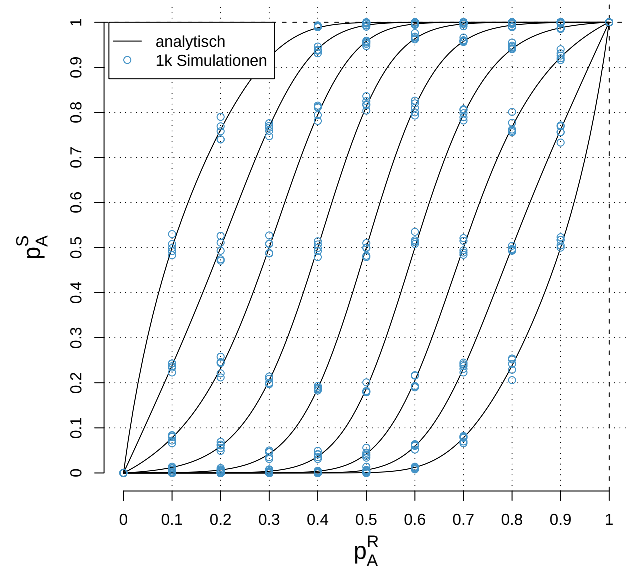
Simulation von Spielen



- Immer bessere Anpassung mit zunehmendem n
- Fehler der MC–Schätzung proportional zu $n^{-1/2}$
- Vergrößerung von n um Faktor 100 $\Rightarrow$ Reduktion Fehler um Faktor $1/10$

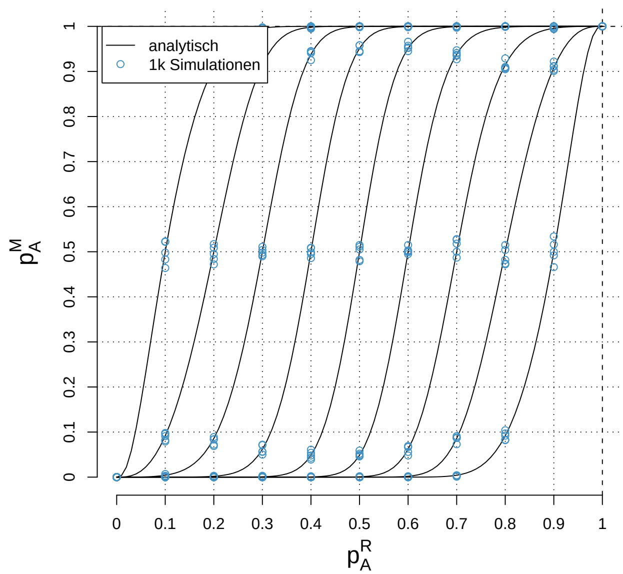
Vergleich Simulation vs. Theorie

Simulation von Sätzen



Vergleich Simulation vs. Theorie

Simulation von Matches



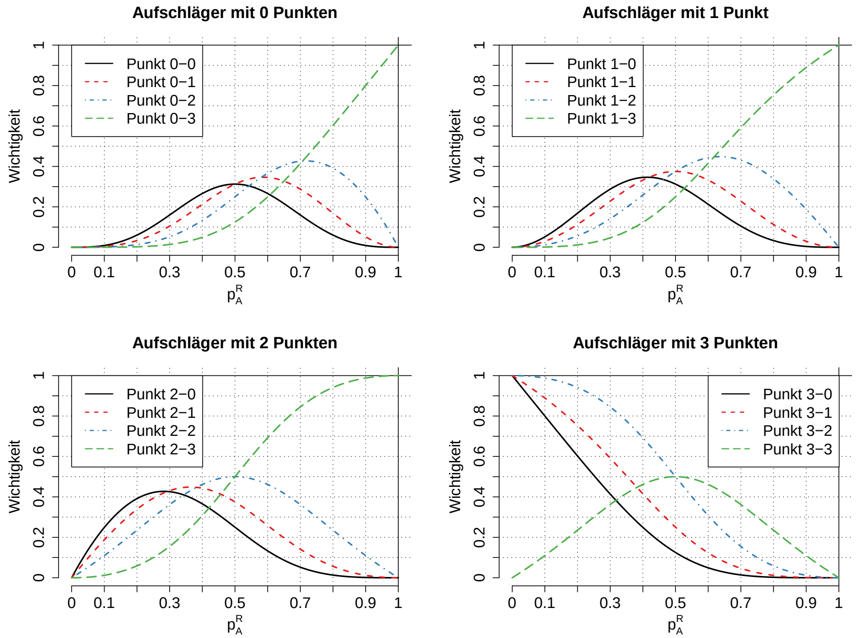
Wichtigkeit von Punkten

- Nachzubildender Effekt: Spieler kann sein Spielvermögen an die Wichtigkeit eines Punktes, Spiels bzw. Satzes anpassen
- Wichtigkeit wird gemessen im Bezug auf den Ausgang eines Spiels, Satzes oder Matches:

$I_{ij}^P = P_{i+1,j}^G - P_{i,j+1}^G$	Wichtigkeit Punkt für Spielgewinn
$I_{ij}^G = P_{i+1,j}^S - P_{i,j+1}^S$	Wichtigkeit Spiel für Satzgewinn
$I_{ij}^T = P_{i+1,j}^T - P_{i,j+1}^T$	Wichtigkeit Punkt für Tiebreakgewinn
$I_{ij}^S = P_{i+1,j}^M - P_{i,j+1}^M$	Wichtigkeit Satz für Matchgewinn

Wichtigkeit von Punkten

Graphishe Darstellung



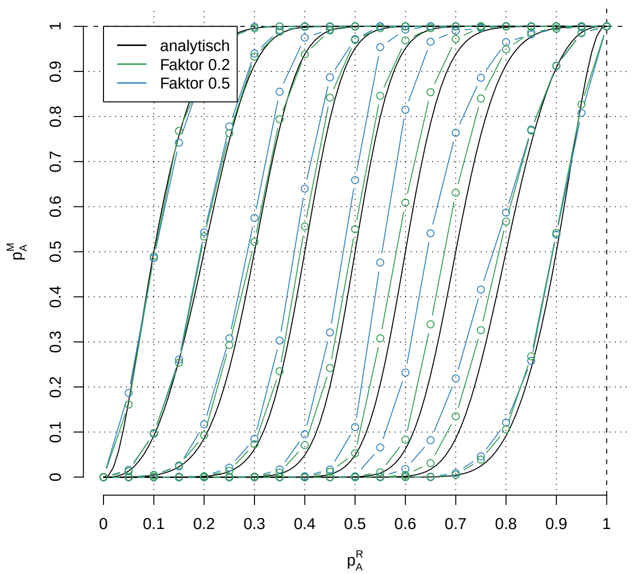
Wichtigkeit von Punkten

Implementierung

- p_A^R wird beim Ausspielen des 'wichtigsten Punktes' um x% erhöht, also
  - für $p_A^R \leq 0.5$ beim Punktestand $3:2$ und bei 'Vorteil' Spieler A
  - für $p_A^R \geq 0.5$ beim Punktestand $2:3$ und bei 'Vorteil' Spieler B
- Beim Ausspielen des am wenigsten wichtigen Punktes wird p_A^R um x% reduziert, also
  - für $p_A^R \leq 0.5$ beim Spielstand $0:3$
  - für $p_A^R \geq 0.5$ beim Spielstand $3:0$

Wichtigkeit von Punkten

Simulation des WvP-Effektes



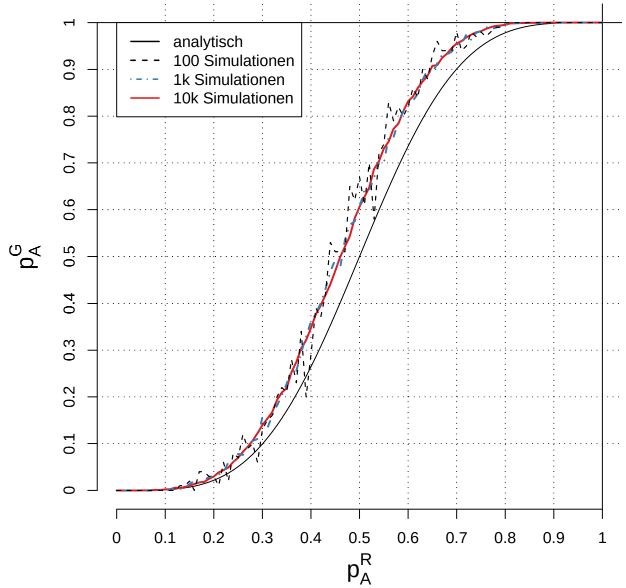
'Hot–Hand'–Effekt

Definition und Implementierung

- Nachzubildender Effekt: Spieler spielt umso besser, je erfolgreicher er kurzfristig ist, bzw. wenn er einen 'Lauf' hat
- Modellierung: Punktgewinnwahrschienlichkeit p_A^R wird für den einem gewonnenen Punkt unmittelbar nachfolgenden Punkt um x% erhöht.

'Hot–Hand'–Effekt

Wirkung des Effektes auf p_A^G



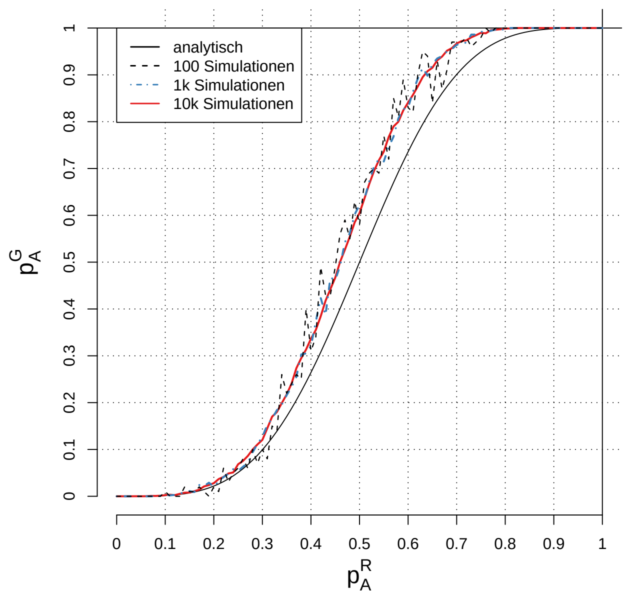
'Back–to–the–Wall'–Effekt

Definition und Implementierung

- Dem 'Hot–Hand'–Effekt entgegengesetz (bezüglich des Ausgangspunktes, nicht der Wirkung)
- Nachzubildender Effekt: Spieler spielt besser, wenn er am verlieren ist, mit dem 'Rücken zur Wand' steht
- Modellierung: p_A^R wird für den einem verlorenen Punkt unmittelbar nachfolgenden Ballwechsel um x% erhöht

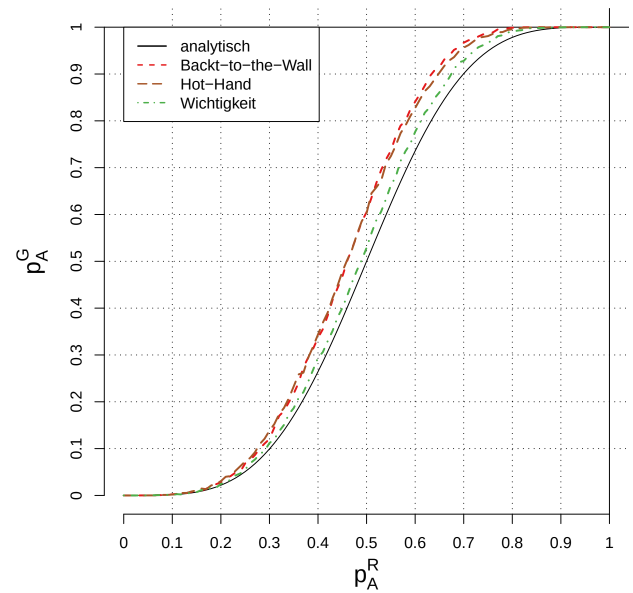
'Back–to–the–Wall'–Effekt

Wirkung des Effektes auf p_A^G



Modelle im Vergleich

'Hot–Hand' vs. 'Back–to–the–Wall'



Fazit

- Zählweise beim Tennis begünstigt den besseren Spieler
- Monte–Carlo–Verfahren geeignet um Tennisspiele, -sätze und -matches nachzubilden
- MC–Schätzer liefern für hinreichend großes n gute Näherungen an die Theorie
- MC–Methode geeignet zum Modellieren von Effekten mit abhängigen Punkten

Ausblick

Weitere Anwendungsmöglichkeiten

- Prädiktive Rankings (im Ggs. zu ausgansbasierten)
- Simulation von Turnieren (auf Basis der p^R vorheriger Runden)
- Modellierung weiterer Effekte mit Abhängigkeiten
- Prüfen der Auswirkungen verschiedener Zählsysteme, unterschiedliche Turnierregeln, u. v. a.
- Einfache Übertragung auf verwandte Sportarten (z. B. Tischtennis, Badminton)