

Pythagoras and the



- ❖ Die Sportart Eishockey
- ❖ Problemstellung im Paper
- ❖ Die Pythagoras Methode
- ❖ Umsetzung im Paper
- ❖ Diskussion der Methodik
- ❖ Übertragung auf DEL
- ❖ Zukunftsaussichten
- ❖ Zusammenfassung

04.12.09

Manuela Bitterlich

2

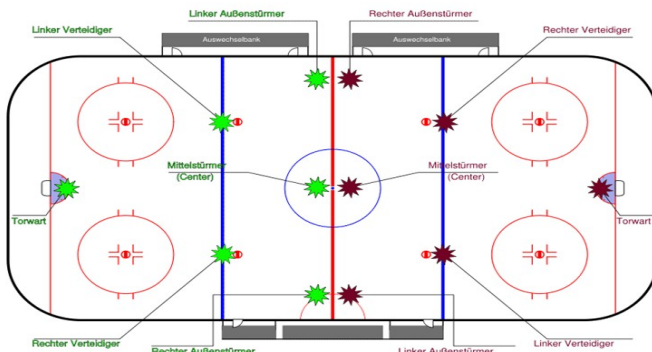
Die Sportart Eishockey



National Hockey League

Eishockey
Problemstellung
Methodik
Umsetzung
Diskussion
DEL
Zukunft

- ❖ Schnellste Mannschaftssportart der Welt
- ❖ Hockey bedeutet „krummer Stock“
- ❖ Entstehung in Kanada



Manuela Bitterlich

04.12.09

3

Eishockey

Problemstellung

Methodik

Umsetzung

Diskussion

DEL

Zukunft

04.12.09

Die Teams in der NHL



- ❖ spielerisch beste Eishockeyliga
- ❖ kleineres Spielfeld



4

Problemstellung im Paper

- Eishockey
- Problemstellung
- Methodik
- Umsetzung
- Diskussion
- DEL
- Zukunft

❖ **Baseball:**
 Untersuchung auf Beziehung zwischen tatsächlichen Gewinnanteil und Anzahl der geschossenen Tore bzw. Anzahl der Gegentore

❖ **Eishockey:**
 Besteht ähnliche Beziehung für unentschiedene Spiele?

$$\text{Gewinn/Verlustanteil} \approx \frac{\text{Siege} + 0.5 (\text{Unentschieden})}{\text{Anzahl aller Spiele}}$$



- Eishockey
- Problemstellung
- Methodik
- Umsetzung
- Diskussion
- DEL
- Zukunft

Die Pythagoras Methode

$$\text{Gewinn-/Verlustanteil} \approx \frac{\text{geschossene Tore}^2}{\text{geschossene Tore}^2 + \text{Gegentore}^2}$$

Die Pythagoras Methode

- Eishockey
- Problemstellung
- Methodik
- Umsetzung
- Diskussion
- DEL
- Zukunft

❖ **Optimierungsproblem:**

- ❖ Finden des optimalen Parameters
- ❖ Minimierungsproblem

❖ **Der mittlere absolute Fehler:**

- ❖ Optimierungsgröße
- ❖ Formel: $\frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}|$



- Eishockey
- Problemstellung
- Methodik
- Umsetzung
- Diskussion
- DEL
- Zukunft

Exponent, der mittleren absoluten Fehler minimiert

- ❖ 1205 Daten aus allen NHL Saisonen bis 2007
- ❖ **Optimierung:**

$$\min \sum_{t=1917}^{2007} \sum_{i=1}^{n_t} \left| wlpct_{it} - \frac{s_{it}^x}{s_{it}^x + a_{it}^x} \right| g_{it}$$

➔ optimaler Exponent, wenn $x = 1.927$

- ❖ Mittlerer absoluter Fehler ca. 2.831 % oder 2.321 Spiele pro Team über 82 NHL Spiele
- ❖ Exponenten, bei Saisonen in Klassen eingeteilt, unterscheiden sich kaum

Drei eindeutige Exponenten

- ❖ Andere Ansätze gewähren variable Exponenten
- ❖ Zunächst drei eindeutige Exponenten

$$\min \sum_{t=1917}^{2007} \sum_{i=1}^{n_t} \left| wlpct_{it} - \frac{s_{it}^{x_1}}{s_{it}^{x_2} + a_{it}^{x_3}} \right| g_{it}$$

- ❖ Optimale Werte:
 $x_1 = 2.045$, $x_2 = 2.0526$ und $x_3 = 2.0306$
- ❖ Mittlerer absoluter Fehler ca. 2.7769 % oder 2.277 Spiele pro Team über 82 NHL Spiele

04.12.09

Manuela Bitterlich

9

Beispiel anhand der Toronto Maple Leafs

- ❖ 19. Tabellenplatz in der Saison 2006/2007
- ❖ 254 Tore, 262 Gegentore, 32 Siege, 31 Verluste und 19 Unentschieden
- ❖ Tatsächlicher Gewinnanteil: 50.61 %
- eindeutiger Exponent → 48.51 %
- drei eindeutige Exponenten → 49.35 %
- Modell mit Interaktionen → 51.05 %

04.12.09

Manuela Bitterlich

11

Einbeziehung von Interaktionen

- ❖ Regressionsansatz mit Interaktionen:

$$\min \sum_{t=1917}^{2007} \sum_{i=1}^{n_t} \left| wlpct_{it} - \left(x_0 + \frac{x_1 + x_2 s_{ikt}^{x_3} + x_4 a_{ikt}^{x_5} + x_6 (s_{ikt} a_{ikt})^{x_7}}{x_8 + x_9 s_{ikt}^{x_{10}} + x_{11} a_{ikt}^{x_{12}} + x_{13} (s_{ikt} a_{ikt})^{x_{14}}} \right) \right| g_{it}$$

- ❖ Mit einsetzen der optimalen Werte erhält man:

$$E[wlpct_{it}] = 0.0002 + \frac{0.798 s_{it}^{2.15}}{2.518 s_{it}^{1.96} + 0.041 a_{it}^{2.62} + 0.455 (s_{it} a_{it})^{0.92}}$$

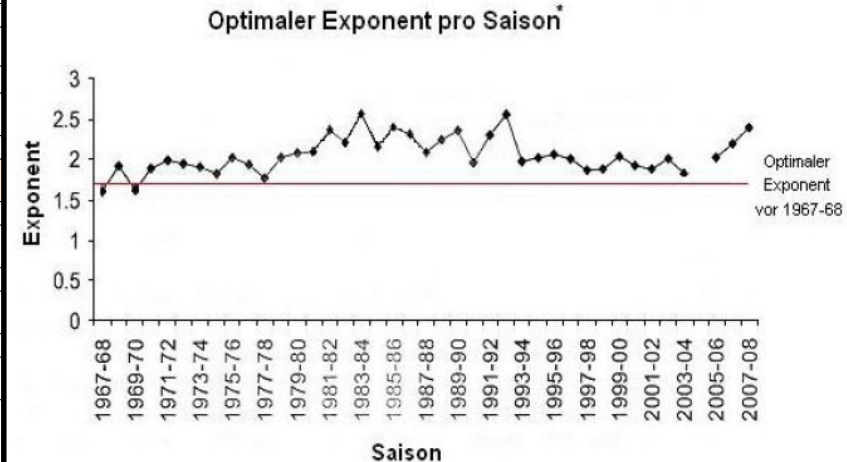
- ❖ Mittlerer absoluter Fehler ca. 2.675 % oder 2.193 Spiele pro Team über 82 NHL Spiele

04.12.09

Manuela Bitterlich

10

Entwicklung des Exponenten über die Saisonen



* Die NHL Saison 2004-05 wurde gestrichen wegen Problemen mit der Lizenzvergabe

04.12.09

Manuela Bitterlich

12

Lösung im Paper

Eishockey
Problemstellung
Methodik
Umsetzung
Diskussion
DEL
Zukunft

- ❖ Einfachste Formel:

$$\text{Gewinn-/Verlustanteil} \approx \frac{\text{geschossene Tore}^x}{\text{geschossene Tore}^x + \text{Zahl der Gegentore}^x}$$

Übertragung auf DEL

Eishockey
Problemstellung
Methodik
Umsetzung
Diskussion
DEL
Zukunft

- ❖ Optimierung in R: `optim()`
- ❖ Optimale Exponenten der letzten drei Saisonen

2008/09: $x = 1.683877$	}	→	$x = 2.073364$
2007/08: $x = 2.318670$			
2006/07: $x = 1.961608$			
- berechnet über Formel für Unentschieden
- ❖ ähnliche Streuung bei Formel für Siege
- ❖ Begründet durch Streuung der Tordifferenzen
- ❖ Von 2000-2009: $x = 2.027383$

Diskussion der Methodik

Eishockey
Problemstellung
Methodik
Umsetzung
Diskussion
DEL
Zukunft

- ❖ Einfache Formel, für alle gut zu verstehen
- ❖ Komplexe Formen: zunehmende Genauigkeit, aber Vorteil zu gering
- ❖ Interessanter Aspekt, dass über Anzahl an Tore und Anzahl an Gegentore auf Gewinnwahrscheinlichkeit zu schließen ist
- ❖ Kein Vergleich mit den anderen genannten Ansätzen
- ❖ Methodik beruht auf Vergangenheitsdaten

Zukunftsaussichten für weitere Analysen

Eishockey
Problemstellung
Methodik
Umsetzung
Diskussion
DEL
Zukunft

- ❖ individuelle Spielerstatistiken
- ❖ Optimierung auf weitere Klassen von funktionellen Formeln ausweiten
- ❖ auf alle Eishockey-Ligen anwendbar
- ❖ Gut auf andere Sportarten übertragbar
- ❖ beste bereits bestehende Statistik: Plus/Minus-Statistik



Zusammenfassung

Eishockey

❖ Einfachste Formel:

Problemstellung

$$\text{Gewinn-/Verlustanteil} \approx \frac{\text{geschossene Tore}^x}{\text{geschossene Tore}^x + \text{Zahl der Gegentore}^x}$$

Methodik

❖ Komplexität liefert nur geringen Vorteil an Genauigkeit

Umsetzung

Diskussion

❖ Fast kein Unterschied zwischen Exponent der NHL und der DEL

DEL

Zukunft

❖ beste Statistik im Eishockey: +/- Statistik