Ein neues Verfahren zur Abrechnung von Teamturnieren im Bridge

Gerben Dirksen und Michael Stieglitz

5. Juli 2005*

Zusammenfassung

Im Bridgesport fehlt bei Teamturnieren oft die Zeit dafür, dass jedes Team gegen jedes Team spielt. Um dennoch am Schluss des Turniers eine Rangfolge zu erhalten, kann man ein Verfahren nutzbar machen, das E.Zermelo 1928 für das analoge Problem bei Schachturnieren angegeben hat. Es handelt sich dabei um einen rekursiven Algorithmus zur Berechnung eines Maximum-Likelihood-Schätzers für die Rangfolge. Der Algorithmus lässt sich mit einem einfachen Programm auf jedem Computer durchführen. Nun weiß man aus statistischen Untersuchungen, dass die bei der Abrechnung zwischen zwei Teams auftretenden sogenannten Internationalen Match Points normalverteilt sind mit einer bekannten und von den Teams unabhängigen konstanten Varianz sowie einem von den beiden Teams anhängenden Erwartungswert. Es ist der Inhalt dieser Arbeit, ein Schätzverfahren für den Erwartungswert vorzuschlagen für Teams, die nicht gegeneinander gespielt haben. Unter einer angegeben Internetadresse kann der C++ Quellcode und eine DOS-Version für das gesamte Abrechnungsverfahren abgerufen werden.

Abstract

In team tournaments in bridge there often is not enough time for each team to compete with each other. In order to obtain nevertheless a final order of rank one can apply a procedure proposed by E.Zermelo in 1928 for chess tournaments. This algorithm may be easily implemented on each PC. One knows from tests that the so-called International Match Points occuring during rank computation are normally distributed with constant variance whereas the expectation depends on each pair of teams competing with each other. We propose an algorithm estimating this expectation also for each couple of teams which have not met during the tournament. We also supply an internet address from where the C++ code and a DOS version of the complete algorithm may be downloaded.

*Institut für Astronomie & Astrophysik, Abt. Computational Physics, Universität Tübingen, Auf der Morgenstelle 10, D-72076 Tübingen, Germany, Email dirksen@tat.physik.uni-tuebingen.de und Institut für Mathematische Stochastik, Universität Karlsruhe, Englerstr. 2, D-76128 Karlsruhe, Germany, Email michael.stieglitz@t-online.de

Keywords: ranking in bridge tournaments
MS 2000 Classification: 91A99

1 Das Problem

An einem Turnier beteiligen sich N Teams $T_i, 1 \leq i \leq N$. Spielt jedes Team gegen jedes Team, stimmt also die Menge

$$G : \{(i,j) : 1 \leq i,j \leq N, T_i \text{ spielt gegen } T_j\}$$

der "realen" Begegnungen mit der Menge

$$\overline{G} := \{(i,j) : 1 \leq i,j \leq N, i \neq j\}$$

aller Begegnungen überein, so liest man aus der IMP-VP-Tabelle für die entsprechende Anzahl gespielter Boards die "wahren" Siegpunkte $v_{ij}^*, (i,j) \in \overline{G}$, ab, bildet die Gesamtzahl $v_i^* := \sum_{i \neq j=1}^{N} v_{ij}^*$ der Siegpunkte für Team T_i und ordnet die $(v_i^*)_{1 \leq i \leq N}$ der Größe nach um, um die "wahre" Rangfolge $s^* = (s_i^*)_{1 \leq i \leq N}$ zu erhalten.

Ein Problem entsteht, wenn etwa in einem Clubturnier die Zeit für den Modus "Jeder gegen Jeden" (Round Robin) nicht ausreicht. Dann gibt es Teams, die nicht gegeneinander spielen, d.h. es ist die Menge $G^c := \overline{G} \backslash G$ der "virtuellen" Begegnungen nicht leer, und es muss ein Verfahren bereit gestellt werden, das berücksichtigt, dass nicht alle Teams gegen gleich starke Gegner gespielt haben. Zwar gibt es mehrere solcher Verfahren (z.B. das Schweizer System), doch haben sie in der Regel den gravierenden Nachteil, dass ihre Anwendung Zeit beansprucht und somit die Spielzeit verkürzt. Eine mögliche Alternative bietet das sogenannte Zermelo-Verfahren. Es berechnet eine "faire" Rangfolge für ein nicht vollendetes Round Robin Turnier, indem die nicht gespielten Begegnungen durch die "wahrscheinlichsten" Scores geschätzt werden (vgl. Abschnitt 2).

2 Das Zermelo-Verfahren

Der folgende Vorschlag von M.Stieglitz, bei dem man nach Laufkarten und ohne Zwischenabrechnung wechselt, wird im Bridgeclub Böblingen-Sindelfingen seit 1987 praktiziert: Ausgangspunkt sind die Siegpunkte v_{ij} aus der IMP-VP-Tabelle für $(i,j) \in G$, also für diejenigen Teams T_i, T_j, die gegeneinander gespielt haben. Dann wird nach E.Zermelo (1928) eine Rangliste $s = (s_i)_{1 \leq i \leq N}$ der Spielstärken erstellt, die übrigens im Fall $G = \overline{G}$ die wahre Rangfolge s^* ergibt. Das von Zermelo vorgeschlagene Verfahren ist für den Fall von Schachturnieren konstruiert. In der Modifikation für Bridgeturniere geht man davon aus, dass aus der zunächst unbekannten Rangfolge s die Wahrscheinlichkeiten

$$p_{ij} := \frac{s_i}{s_i + s_j}, (i,j) \in G, \tag{1}$$

dafür gebildet werden, dass T_i das Team T_j besiegt. Ferner werden die aus der IMP-VP-Tabelle bekannten $v_{ij}, (i,j) \in G$, so interpretiert, als ob T_i gegen T_j die Zahl von $(v_{ij} + v_{ji})$ voneinander unabhängigen Kämpfe gespielt hätte, in denen jeweils 1 Siegpunkt mit Wahrscheinlichkeit p_{ij} aus (**??**) zu gewinnen war. Die Wahrscheinlichkeit dafür, dass bei diesen Kämpfen das Team T_i die Zahl von v_{ij} Siegpunkten erhält, ist dann $Bin(v_{ij} + v_{ij}, p_{ij}, v_{ij})$. Dabei ist

$$k \mapsto Bin(n,p,k) := \binom{n}{k} p^k (1-p)^{n-k}$$

die diskrete Dichte von n unabhängigen $\{0,1\}$-verteilten Zufallsvariablen mit Erfolgswahrscheinlichkeit p für 1 Siegpunkt. Das Zermelo-Verfahren berechnet nun iterativ und mit beliebiger Genauigkeit das eindeutig bestimmte s so, dass die über (**??**) von s abhängige Wahrscheinlichkeit

$$\prod_{(i,j) \in G} Bin(v_{ij} + v_{ij}, p_{ij}, v_{ij})$$

maximal wird, d.h. s ein sogenannter Maximum-Likelihood-Schätzer ist (vgl. Abschnitt 4). Dabei wird für $(i,j) \in G^c$ formal $v_{ij} := 0$ gesetzt.

Das Zermelo-Verfahren hängt entscheidend von der IMP-VP-Tabelle ab. Diese aber hat den Nachteil, dass die Aufteilung der VPs unsymmetrisch ist, und zwar zulasten guter Teams, da bei hohen IMP-Werten die Anzahl der VP konstant 25 bleibt, während der Gegner noch durchaus VPs erhalten kann. Aus diesem Grund wird hier ein neues Abrechnungsverfahren vorgeschlagen.

3 Eine Modifikation des Zermelo-Verfahrens

Es sei daran erinnert, dass die Verteilungsfunktion einer $N(a, \sigma^2)$(=normal)-verteilten Zufallsvariable durch

$$\Phi_{a,\sigma^2}(x) := \Phi(\frac{x-a}{\sigma}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-t^2/2} dt = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-a}{\sigma\sqrt{2}}\right)\right], x \in \mathbb{R},$$

definiert ist. Die wesentliche Idee für das Folgende ist die Beobachtung, dass die IMPs K_{ij}, die T_i nach Kampf gegen T_j erhält, $N(I_{ij}, \sigma^2)$-verteilte Zufallsvariablen sind mit von (i,j) unabhängiger Varianz σ^2 und von (i,j) abhängigem Erwartungswert I_{ij}. Da die IMPs K_{ij} nur ganzzahlig sind, muss zunächst geklärt werden, wie diese Aussage zu verstehen ist. Dazu wird als diskrete Dichte κ_{ij} von K_{ij}

$$\kappa_{ij}(k) := \Phi_{I_{ij},\sigma^2}(k + \frac{1}{2}) - \Phi_{I_{ij},\sigma^2}(k - \frac{1}{2}), k \in \mathbb{Z}, (i,j) \in \overline{G}, \tag{2}$$

definiert, d.h. man ordnet den Punkten k diejenige Masse zu, welche die $N(I_{ij}, \sigma^2)$-Verteilung auf dem um k symmetrischen Intervall $[k-\frac{1}{2}, k+\frac{1}{2}]$ hat. Die Varianz σ^2 ist dabei die empirische Varianz der IMPs aus allen gespielten Begegnungen des Turniers. (Statistische Auswertungen wirklich gespielter Teamturniere haben gezeigt, dass die Varianz der IMPs pro gespieltem Board etwa 5.5^2 ist.) Für die realen Begegnungen $(i,j) \in G$ ist I_{ij} aus der IMP-VP-Tabelle bekannt, also κ_{ij} für $(i,j) \in G$ bekannt. Und für die virtuellen Begegnungen $(i,j) \in G^c$ werden die I_{ij} als geschätzte Werte durch das Zermelo-Verfahrens aus den Werten

$$W_{ij} := \begin{cases} \Phi_{0,\sigma^2}(I_{ij}), & (i,j) \in G \\ 0 & (i,j) \in G^c \end{cases} \tag{3}$$

berechnet. Als Ergebnis erhält man die Rangfolge $(S_i)_{1 \le i \le N}$. Dann wird

$$P_{ij} := \frac{S_i}{S_i + S_j}, (i,j) \in G^c,$$

für die virtuellen Begegnungen gesetzt und

$$I_{ij} := \Phi^{-1}_{0,\sigma^2}(P_{ij}), (i,j) \in G^c,$$

definiert. Damit ist κ_{ij} aus (**??**) auch für $(i,j) \in G^c$ definiert. Um die letztlich interessierenden Siegpunkte $V_{ij}, (i,j) \in \overline{G}$, zu definieren, betrachten wir die IMP-VP-Tabelle bei gegebener Anzahl gespielter Boards, die jeder IMP-Differenz $I \in \mathbb{Z}$ die Siegpunkte $g(I)$ mit Werten zwischen 0 und 25 zuordnet. Dann wird

$$V_{ij} := \begin{cases} g(I_{ij}), & (i,j) \in G \\ \sum_{k=0}^{25} g(k)\kappa_{ij}(k), & (i,j) \in G^c \end{cases}$$

gesetzt. D.h. V_{ij} ist der Erwartungswert von $g \circ K_{ij}$, wobei K_{ij} im Fall $(i,j) \in G$ konstant $= I_{ij}$ und im Fall $(i,j) \in G^c$ eine Zufallsvariable mit der diskreten Dichte κ_{ij} aus (**??**) ist. Schließlich erhält man durch

$$V_i := \sum_{i \neq j=1}^{N} V_{ij}, 1 \le i \le N,$$

die Siegpunkte von T_i. Die Rangliste wird dann entsprechend den $(V_i)_{1 \le i \le N}$ gebildet.

4 Der Algorithmus des Zermelo-Verfahrens

Das Verfahren zur Berechnung der Rangfolge $(S_i)_{1 \le i \le N}$ benutzt die Werte $(W_{ij})_{(ij) \in \overline{G}}$ aus (**??**). Gestartet wird die Iteration mit $S_{1i} = 1, 1 \le i \le N$. Definiert man dann

$$S_{k+1,i} := \frac{\sum_{i \neq j=1}^{N} W_{ij}}{\sum_{i \neq j=1}^{N} \frac{W_{ij}+W_{ji}}{S_{ki}+S_{kj}}}, k \in \mathbb{N}, 1 \le i \le N,$$

so ist

$$S_i := \lim_{k \to \infty} S_{ki}, 1 \le i \le N.$$

5 Software

Der C++ Quellcode sowie eine DOS-Version eines Programms für die Zermelo-Auswertung ist kostenlos auf folgender Internetseite herunterzuladen:
http://www.geocities.com/ gerben47/zermelo. Eine Beispiel-Eingabedatei ist beigefügt.

Herrn W.Umlauft danken wir für Hinweise und anregende Diskussionen sowie die Möglichkeit, die neue Methode im Bridgeclub Böblingen-Sindelfingen anzuwenden.

References

Zermelo, E. (1928), Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung, Math.Z.29, 436-460.