E. Fabri *giugno 2015*

Sul punteggio SB

Introduzione

Voglio qui dare una spiegazione in termini semplici di una relazione che vale per i punteggi SB nei tornei di scacchi. Sarà anche una risposta alla domanda fatta di recente da ilMusso in it.hobby.scacchi.(1) Scrive ilMusso:

Come sappiamo, il sistema Sonneborn-Berger prevede che per calcolare tale punteggio per ciascun giocatore venga assegnato il totale dei punti ottenuti da ogni avversario battuto nello scontro diretto, la metà dei punti ottenuti da ogni avversario con cui si è avuto un risultato di pareggio e zero dei punti ottenuti dagli avversari dai quali si è stati battuti.

Facendo la somma dei punteggi SB dei giocatori partecipanti al torneo con girone all'italiana di Chanty-Mansijsk conclusosi qualche giorno fa si ottiene un totale di 353,00.

Concludeva chiedendo il perché del risultato, e se esiste una formula. La risposta di Paolo Casacchi è stata questa:

Somma di tutti gli SB $= N(N-1)^2/2 - \sum_j P_j^2$ *dove* P_j *è il punteggio del* j*-mo giocatore*

ma è stata giudicata troppo astrusa (la formula e ancor più la deduzione).

Riprenderò quindi la questione daccapo, in modo — spero — più comprensibile, e ritroverò la formula di Casacchi.

Notazioni

Indico i giocatori con lettere maiuscole: A, B, ... oppure minuscole quando li uso come indici. R_{ab} è il risultato della partita tra A e B, *visto da* A; quindi vale 1 se A ha vinto, 0 se ha perso, 0.5 in caso di patta. Per la stessa partita, R_{ba} è il risultato *visto da* B; quindi $R_{\mathrm{ba}} = 0$ se $R_{\mathrm{ab}} = 1$, ecc. Notiamo che è sempre

$$R_{\mathrm{ab}} + R_{\mathrm{ba}} = 1. \tag{1}$$

Per rappresentare l'intero torneo conviene usare una *matrice*. In parole semplici, una matrice non è che una scacchiera dove in ogni casella è scritto un numero. Solo che invece di essere 8×8, la scacchiera avrà tante traverse ("righe" quando si parla di matrici) e tante colonne quanti sono i giocatori: 12 nell'esempio fatto da ilMusso, N in generale. A differenza della notazione scacchistica, in cui le colonne sono indicate da lettere e le traverse da numeri,

(1) `http://groups.google.com/d/msg/it.hobby.scacchi/Hq_Aw7pqGy8/3TdWkkufuiAJ`

qui userò lettere in entrambi i casi. Quindi quella che negli scacchi sarebbe "b3," la indicherò con "bc." Vedremo subito che c'è un vantaggio a far questo.

Ogni casella della matrice corrisponde a una partita: per es. "bf" corrisponde alla partita tra B ed F. Ma anche "fb" corrisponde alla stessa partita; che motivo c'è di averla due volte? Il motivo è che in tal modo possiamo vederla una volta dal punto di vista di B, e un'altra dal punto di vista di F. Chiariamo meglio.

Dicevo che nelle caselle mettiamo dei numeri: più precisamente, ci mettiamo i *risultati* della partita, visti dai due giocatori. Così se B ha vinto su F, in "bf" metteremo $R_{\mathrm{bf}} = 1$, in "fb" metteremo $R_{\mathrm{fb}} = 0$. La (1) ci dice che la somma dei valori nelle due caselle è sempre 1. Ma dove si trovano queste due caselle? È facile capire che sono simmetriche rispetto alla diagonale ascendente che parte dalla casella "aa."

C'è da tener conto di un'eccezione: le caselle nella diagonale ascendente di cui sopra, ossia "aa," "bb," ecc. Queste non corrispondono a nessuna partita (un giocatore non gioca contro se stesso) quindi andrebbero lasciate vuote. Converrà però scriverci il valore 0, ossia definire $R_{\mathrm{aa}} = 0$, $R_{\mathrm{bb}} = 0$, ecc.

Calcolo dei punteggi

In primo luogo consideriamo il punteggio che chiamerò "semplice," dato che non conosco il nome ufficiale (non sono uno scacchista). Lo indico con P_{a} per il giocatore A, ecc. Consiste nella somma dei punteggi ottenuti da ciascun giocatore nelle $N - 1$ partite da lui disputate. Così ad es. avremo

$$P_{\mathrm{a}} = R_{\mathrm{ab}} + R_{\mathrm{ac}} + \cdots \tag{2}$$

$$P_{\mathrm{a}} = R_{\mathrm{aa}} + R_{\mathrm{ab}} + R_{\mathrm{ac}} + \cdots \tag{2'}$$

che è la somma di tutti i numeri della colonna "a" nella matrice. (Ecco perché abbiamo posto $R_{\mathrm{aa}} = 0\,!$). Idem per gli altri giocatori e le altre colonne.

Il punteggio SB (Sonneborn–Berger) di A è definito come segue:

$$S_{\mathrm{a}} = R_{\mathrm{ab}}P_{\mathrm{b}} + R_{\mathrm{ac}}P_{\mathrm{c}} + \cdots \tag{3}$$

Infatti se per es. A ha vinto su B prende l'intero punteggio di B; se ha pareggiato, ne prende metà, perché $R_{\mathrm{ab}} = 0.5$; se ha perso non prende niente. Invece di (3) posso anche scrivere

$$S_{\mathrm{a}} = R_{\mathrm{aa}}P_{\mathrm{a}} + R_{\mathrm{ab}}P_{\mathrm{b}} + R_{\mathrm{ac}}P_{\mathrm{c}} + \cdots \tag{3'}$$

per la solita ragione.

La somma dei punteggi

Non resta ora che fare la somma di tutte le S_a, S_b, …: chiamiamola T. Abbiamo

$$\begin{aligned}T = &(R_\mathrm{aa}P_\mathrm{a} + R_\mathrm{ab}P_\mathrm{b} + R_\mathrm{ac}P_\mathrm{c} + \cdots) + \\ &(R_\mathrm{ba}P_\mathrm{a} + R_\mathrm{bb}P_\mathrm{b} + R_\mathrm{bc}P_\mathrm{c} + \cdots) + \\ &(R_\mathrm{ca}P_\mathrm{a} + R_\mathrm{cb}P_\mathrm{b} + R_\mathrm{cc}P_\mathrm{c} + \cdots) + \\ &\ldots\end{aligned}$$

Riordinando e raccogliendo P_a, P_b, … si ottiene

$$\begin{aligned}T = &(R_\mathrm{aa} + R_\mathrm{ba} + R_\mathrm{ca} + \cdots)P_\mathrm{a} + \\ &(R_\mathrm{ab} + R_\mathrm{bb} + R_\mathrm{cb} + \cdots)P_\mathrm{b} + \\ &(R_\mathrm{ac} + R_\mathrm{bc} + R_\mathrm{cc} + \cdots)P_\mathrm{c} + \\ &\ldots\end{aligned} \tag{4}$$

Guardiamo ora l'espressione in parentesi nella prima riga:

$$R_\mathrm{aa} + R_\mathrm{ba} + R_\mathrm{ca} + \cdots$$

Usando la (1) possiamo scriverla

$$R_\mathrm{aa} + (1 - R_\mathrm{ab}) + (1 - R_\mathrm{ac}) + \cdots = (N - 1) - P_\mathrm{a}. \tag{5}$$

Questa però va spiegata…

Quanti sono gli 1 a primo membro? Sono tanti quanti i termini della somma, escluso il primo; quindi $N - 1$. Se poi ricordiamo che $R_\mathrm{aa} = 0$, vediamo che tutti gli altri termini sommati danno proprio $-P_\mathrm{a}$.

Quello che la (5) dice per la prima riga della (4), vale anche per tutte le altre righe. Quindi la (4) si può trasformare in

$$\begin{aligned}T = &[(N - 1) - P_\mathrm{a}]\,P_\mathrm{a} + [(N - 1] - P_\mathrm{b}]\,P_\mathrm{b} + [(N - 1) - P_\mathrm{c}]\,P_\mathrm{c} + \cdots = \\ &(N - 1)(P_\mathrm{a} + P_\mathrm{b} + P_\mathrm{c} + \cdots) - (P_\mathrm{a}^2 + P_\mathrm{b}^2 + P_\mathrm{c}^2 + \cdots).\end{aligned}$$

Resta solo da calcolare $P_\mathrm{a} + P_\mathrm{b} + P_\mathrm{c} + \cdots$: la somma dei tutti i punteggi semplici. Ma questo è facile, perché ogni partita contribuisce per 1 alla somma. Per es. la partita "ab" dà un punto ad A se è lui che l'ha vinta (e 0 a B). Se invece ha vinto B, succede l'inverso. Se infine è stata patta, tocca mezzo punto ciascuno, ma sempre un punto in tutto fra i due.

Basta ricordare che il numero di partite è $\frac{1}{2}N(N - 1)$ per arrivare alla conclusione:

$$P_\mathrm{a} + P_\mathrm{b} + P_\mathrm{c} + \cdots = \tfrac{1}{2}N(N - 1) \tag{6}$$

e infine

$$T = \tfrac{1}{2}N(N-1)^2 - (P_\mathrm{a}^2 + P_\mathrm{b}^2 + P_\mathrm{c}^2 + \cdots) \tag{7}$$

che è la formula di Casacchi.

Qualche commento

Il secondo membro della (7) consiste di due termini: il primo è fisso, dato il numero di giocatori, mentre il secondo è variabile con i punteggi. Sappiamo dalla (6) che la somma dei punteggi è fissa, ed è noto che la somma dei quadrati di N numeri la cui somma sia fissata, assume il valore minimo quando sono tutti uguali.

Quando i P_x sono tutti uguali, valgono ciascuno $\frac{1}{2}(N-1)$ e la somma dei quadrati fa $\frac{1}{4}N(N-1)^2$. Di conseguenza il *massimo* di T è $\frac{1}{4}N(N-1)^2$. Se $N = 12$, abbiamo $T = 363$.

Come si fa ad avere punteggi tutti uguali? Un modo ovvio è supporre che tutte le partite siano finite patte. Non è l'unico, ma non importa.

È meno facile capire in qual caso la somma dei quadrati dei punteggi sarà massima. A intuito, bisogna che i punteggi siano i più diversi possibili. Supponiamo per es. che A vinca tutte le partite, totalizzando $N-1$ punti, e che un altro giocatore, diciamo Z, le perda tutte, con punteggio 0. I giocatori rimanenti avranno per forza punteggi intermedi: infatti il più scarso avrà vinto almeno con Z, e quindi avrà almeno un punto; il più bravo avrà perso almeno con A, e potrà avere al massimo $N-2$ punti, se avrà vinto con tutti gli altri.

Insomma la distribuzione di punteggi meno uniforme possibile sembra essere $N-1, N-2, \ldots, 1, 0$. La somma dei quadrati è $\frac{1}{6}N(N-1)(2N-1)$ e per T si ottiene $\frac{1}{6}N(N-1)(N-2)$. Se $N = 12$, si ha $T = 220$.

Attenzione: questa non è una dimostrazione! Il minimo di T è solo plausibile, non certo. Il minimo vero *potrebbe* essere ancora minore.

Il risultato 353 del torneo di Chanty-Mansijsk è molto più vicino al massimo che al minimo di T, il che indica che i concorrenti erano molto equilibrati. Anche senza essere scacchista, lo trovo piuttosto ovvio.

Non so se ciò che ho scritto riuscirà comprensibile. Di sicuro è molto più lungo della dimostrazione di Casacchi. È il prezzo che si paga a non voler usare sommatorie e altri artifici matematici...