

# Simulando Resultados para Previsão em Competições Esportivas

José Wilson C. Pinto\*, Robson M. da Silva,

Depto de Matemática, ICE, UFRRJ,  
23890-000, Seropédica, RJ

E-mail: jpinto@ic.uff.br, rmariano@ufrj.br,

Elias M. da Silva, Ewerton C. Martins, João M. L. Jr, Patrick P. Félix

Depto de Informática, FAMERC,  
Rua Mercúrio, 293

21525-660, Rio de Janeiro, RJ

E-mail: e.moraess@ig.com.br, ewertoncmap@uol.com.br, jmarcoslj@bol.com.br, patrick.felix@ig.com.br

Em tempos atuais, é comum observarmos na mídia, matemáticos, estatísticos e comentaristas esportivos, apresentarem em seus noticiários as chances que esta ou aquela equipe possui de sagrar-se campeã, se classificar ou ser rebaixada em campeonatos de futebol, organizados sobre o sistema de pontos corridos. Na maioria das vezes, o método utilizado para se chegar a estes números não são apresentados.

O principal objetivo deste trabalho é apresentar uma proposta para calcular a probabilidade que cada uma das equipes possui de conquistar o título. Desta forma, desenvolveu-se um procedimento que, a partir da tabela de confrontos da competição e dos resultados anteriormente apresentados, são simulados os próximos placares até o final do campeonato.

Calcula-se em determinada rodada, a probabilidade que cada equipe possui de conquistar o título, executando-se em qualquer rodada do campeonato o procedimento **CalculaProb**, onde **M** é o número de iterações do procedimento e **C<sub>i</sub>** o número de iterações em que a equipe **i** sagrou-se campeã, então **P<sub>i</sub> = C<sub>i</sub>/M** é a probabilidade que a equipe **i** possui de ser campeã.

O procedimento **CalculaProb**, tem como parâmetros de entrada **T**, que é a tabela do campeonato com todas as partidas e os placares de todas as já realizadas, **R** que representa a rodada em que se está executando o procedimento e **M** número de iterações do procedimento. Inicialmente, cria-se o vetor **C**, com uma entrada para cada equipe e será responsável por armazenar a quantidade de vezes que cada equipe aparece como campeã. Na linha 3,

**Procedimento CalculaProb(T,M,R);**

```
1 C ← InicializaVetor();
2 Para a = 1 até M faça
3   J ← InicializaJogos(T,R);
4   Para todo ji ∈ J faça
5     Simula(ji);
6   Fim-para;
7   C ← AtualizaVetor(T,J);
8 Fim-para;
9 Retornar(C);
Fim.
```

inicializa-se uma lista **J**, que contém os jogos que pertencem as rodadas posteriores a rodada **R**, sem seus placares. Os placares para estes jogos serão simulados pela função **Simula()**. O vetor **C**, é atualizado na linha 7, de acordo com a

equipe apontada como campeã nesta iteração.

A função **Simula()**, utiliza-se do histórico da equipe em confrontos anteriores, ou seja, sua campanha até a rodada **R**, e uma componente aleatória, de forma que, a imprevisibilidade do futebol seja representada. A simulação do placar de uma partida entre o time da casa (**timeC**) e o time visitante (**timeV**), é realizada da seguinte forma: seja **r**, a rodada a qual pertence a partida a ser simulada. Os pesos dos gols marcados(**GM**) e gols sofridos(**GS**) por cada equipe, são calculados através das seguintes expressões:

$$GM_c = \sum_{i=1}^{r-1} i \cdot gm_i^c \cdot \delta \quad GS_c = \sum_{i=1}^{r-1} i \cdot gs_i^c \cdot \delta$$

$$GM_v = \sum_{i=1}^{r-1} i \cdot gm_i^v \cdot \delta \quad GS_v = \sum_{i=1}^{r-1} i \cdot gs_i^v \cdot \delta$$

onde os índices **c** e **v**, representam time da casa e visitante, respectivamente, e **i** é o número da rodada. Sendo assim, **gm<sub>i</sub><sup>c</sup>** é a quantidade de gols marcados pelo time **c** na rodada **i**, assim como, **gs<sub>i</sub><sup>c</sup>** a quantidade de gols sofridos pelo time **c** na rodada **i**. De forma análoga, são calculados os pesos para o time visitante. A variável **δ=1**, quando na rodada **i** a equipe da casa jogou como visitante e **δ=2**, quando jogou em casa. Para o time visitante, **δ=2** quando na rodada **i** a equipe jogou como visitante e **δ=1**, quando jogou em casa. Calcula-se então os gols do timeC, nesta partida, de acordo com as expressões abaixo. Inicialmente, calcula-se o **PC<sub>c</sub>**, percentual de chances. Em seguida, o **NC<sub>c</sub>**, número de chances de gol, que é

$$PC_c = \left( \frac{GM_c + GS_v}{GM_c + GS_c + GM_v + GS_v} \right)$$

um número inteiro, arredondado pela função **round()**, aplicada no produto entre **PC<sub>c</sub>** e 7. O valor 7 foi calibrado de acordo com a média de gols das competições estudadas.

$$NC_c = \text{round}(PC_c \times 7)$$

$$Gols_c = \sum_{i=1}^{NC_c} b_n \leftarrow \text{rand}\{0,1\}$$

Dessa forma, o número de gols(**Gols<sub>c</sub>**) é calculado através deste somatório onde, **b<sub>n</sub>** é uma variável binária {0,1}, gerada

aleatoriamente pela função **rand{0,1}**.  
As tabelas abaixo, apresentam uma comparação entre o calculo realizado de forma puramente matemática(A%), e o calculo realizado pela função **CalculaProb()**(S%). Para esta comparação foram realizadas 1 milhão de iterações, para cada um dos testes. Estes testes foram realizados em quatro rodadas diferentes do campeonato brasileiro de 2004, permitindo assim, avaliar o desempenho do procedimento.

| 15ª RODADA |    |   |   |    |    |     |       |       | 25ª RODADA |    |    |    |    |    |     |       |       | 35ª RODADA |    |    |    |    |    |     |       |       | 45ª RODADA |    |    |    |     |    |     |       |       |
|------------|----|---|---|----|----|-----|-------|-------|------------|----|----|----|----|----|-----|-------|-------|------------|----|----|----|----|----|-----|-------|-------|------------|----|----|----|-----|----|-----|-------|-------|
| Equipe     | P  | V | E | GP | GC | S   | A%    | S%    | Equipe     | P  | V  | E  | GP | GC | S   | A%    | S%    | Equipe     | P  | V  | E  | GP | GC | S   | A%    | S%    | Equipe     | P  | V  | E  | GP  | GC | S   | A%    | S%    |
| SAN        | 28 | 9 | 1 | 29 | 22 | 7   | 17.8% | 24.2% | PAL        | 46 | 13 | 7  | 41 | 22 | 19  | 22.3% | 24.3% | CAP        | 68 | 20 | 8  | 69 | 38 | 31  | 63.3% | 68.8% | SAN        | 86 | 26 | 8  | 101 | 57 | 44  | 77.8% | 78.5% |
| PAL        | 27 | 7 | 6 | 28 | 12 | 16  | 13.3% | 28.0% | SAN        | 44 | 14 | 2  | 56 | 39 | 17  | 15.3% | 21.9% | SAN        | 65 | 20 | 5  | 78 | 51 | 27  | 29.9% | 28.4% | SPA        | 85 | 25 | 10 | 92  | 55 | 37  | 22.2% | 21.5% |
| FIG        | 27 | 7 | 6 | 21 | 11 | 10  | 13.1% | 12.6% | PON        | 44 | 14 | 2  | 26 | 33 | -7  | 14.7% | 1.5%  | SPA        | 59 | 17 | 8  | 52 | 33 | 19  | 2.8%  | 1.9%  | SPA        | 82 | 24 | 10 | 78  | 41 | 37  | 0.0%  | 0.0%  |
| SPA        | 25 | 7 | 4 | 20 | 16 | 4   | 8.7%  | 4.0%  | SPA        | 44 | 13 | 5  | 34 | 23 | 11  | 13.8% | 20.3% | JUV        | 59 | 17 | 8  | 49 | 43 | 6   | 2.5%  | 0.6%  | PAL        | 78 | 22 | 12 | 49  | 46 | 25  | 0.0%  | 0.0%  |
| CAP        | 24 | 7 | 3 | 25 | 14 | 11  | 7.0%  | 16.4% | JUV        | 44 | 13 | 5  | 33 | 27 | 6   | 13.7% | 10.7% | PAL        | 56 | 15 | 11 | 55 | 36 | 19  | 0.6%  | 0.3%  | COR        | 71 | 19 | 14 | 71  | 52 | -3  | 0.0%  | 0.0%  |
| CRU        | 24 | 7 | 3 | 20 | 21 | -1  | 6.6%  | 1.5%  | CAP        | 42 | 12 | 6  | 42 | 27 | 15  | 7.8%  | 13.8% | GOI        | 55 | 16 | 7  | 64 | 55 | 9   | 0.5%  | 0.1%  | JUV        | 70 | 20 | 10 | 59  | 64 | -5  | 0.0%  | 0.0%  |
| INT        | 23 | 7 | 2 | 19 | 19 | 0   | 5.2%  | 1.8%  | GOI        | 42 | 12 | 6  | 48 | 41 | 7   | 7.6%  | 5.2%  | PON        | 53 | 16 | 5  | 31 | 51 | -20 | 0.1%  | 0.0%  | GOI        | 69 | 20 | 9  | 79  | 68 | 11  | 0.0%  | 0.0%  |
| PON        | 23 | 7 | 2 | 13 | 19 | -6  | 5.2%  | 0.2%  | COR        | 37 | 10 | 7  | 29 | 35 | -6  | 1.3%  | 0.7%  | FLU        | 53 | 14 | 11 | 49 | 48 | 1   | 0.1%  | 0.0%  | FLU        | 66 | 18 | 12 | 64  | 67 | -3  | 0.0%  | 0.0%  |
| GOI        | 23 | 6 | 5 | 27 | 27 | 0   | 4.9%  | 1.0%  | CTB        | 37 | 9  | 10 | 28 | 24 | 4   | 1.2%  | 0.7%  | COR        | 53 | 14 | 11 | 37 | 42 | -5  | 0.1%  | 0.0%  | INT        | 64 | 19 | 7  | 64  | 58 | 6   | 0.0%  | 0.0%  |
| VIT        | 21 | 6 | 3 | 27 | 21 | 6   | 3.1%  | 2.4%  | CRI        | 35 | 10 | 5  | 40 | 38 | 2   | 0.7%  | 0.3%  | CTB        | 49 | 12 | 13 | 39 | 35 | 4   | 0.0%  | 0.0%  | FIG        | 63 | 17 | 12 | 55  | 54 | 1   | 0.0%  | 0.0%  |
| JUV        | 21 | 6 | 3 | 18 | 19 | -1  | 2.9%  | 0.8%  | FIG        | 35 | 8  | 11 | 28 | 25 | 3   | 0.6%  | 0.1%  | FIG        | 48 | 12 | 12 | 44 | 44 | 0   | 0.0%  | 0.0%  | PON        | 61 | 18 | 7  | 41  | 72 | -31 | 0.0%  | 0.0%  |
| FLU        | 21 | 5 | 6 | 18 | 19 | -1  | 2.8%  | 0.6%  | INT        | 34 | 10 | 4  | 35 | 32 | 3   | 0.5%  | 0.2%  | CRU        | 47 | 13 | 8  | 54 | 56 | -2  | 0.0%  | 0.0%  | CTB        | 61 | 15 | 16 | 50  | 45 | 5   | 0.0%  | 0.0%  |
| VAS        | 20 | 5 | 5 | 23 | 18 | 5   | 2.2%  | 3.8%  | CRU        | 34 | 10 | 4  | 32 | 35 | -3  | 0.4%  | 0.1%  | INT        | 43 | 12 | 7  | 44 | 46 | -2  | 0.0%  | 0.0%  | CRU        | 56 | 16 | 8  | 67  | 75 | -8  | 0.0%  | 0.0%  |
| CRI        | 19 | 5 | 4 | 20 | 22 | -2  | 1.6%  | 0.3%  | FLU        | 32 | 8  | 8  | 31 | 35 | -4  | 0.1%  | 0.0%  | PAY        | 43 | 11 | 10 | 44 | 61 | -17 | 0.0%  | 0.0%  | PAR        | 54 | 15 | 9  | 51  | 71 | -20 | 0.0%  | 0.0%  |
| COR        | 19 | 5 | 4 | 16 | 27 | -11 | 1.5%  | 0.4%  | VAS        | 30 | 8  | 6  | 38 | 35 | 3   | 0.1%  | 0.1%  | VAS        | 42 | 11 | 9  | 51 | 50 | -1  | 0.0%  | 0.0%  | VAS        | 54 | 14 | 12 | 63  | 66 | -3  | 0.0%  | 0.0%  |
| CTB        | 18 | 4 | 6 | 15 | 15 | 0   | 1.1%  | 0.7%  | VIT        | 29 | 8  | 5  | 41 | 44 | -3  | 0.0%  | 0.0%  | CRU        | 40 | 11 | 7  | 46 | 55 | -9  | 0.0%  | 0.0%  | SCA        | 53 | 23 | 8  | 65  | 46 | 19  | 0.0%  | 0.0%  |
| PAM        | 17 | 4 | 5 | 14 | 22 | -8  | 0.8%  | 0.4%  | CAM        | 28 | 6  | 10 | 33 | 36 | -3  | 0.0%  | 0.0%  | BOT        | 39 | 8  | 15 | 48 | 54 | -6  | 0.0%  | 0.0%  | PAY        | 53 | 13 | 14 | 54  | 75 | -21 | 0.0%  | 0.0%  |
| CAR        | 17 | 3 | 8 | 18 | 20 | -2  | 0.7%  | 0.5%  | GRE        | 26 | 6  | 8  | 35 | 40 | -5  | 0.0%  | 0.0%  | CAM        | 39 | 8  | 15 | 43 | 52 | -9  | 0.0%  | 0.0%  | FLA        | 51 | 12 | 15 | 45  | 51 | -6  | 0.0%  | 0.0%  |
| GUA        | 16 | 4 | 4 | 13 | 19 | -6  | 0.5%  | 0.1%  | PAY        | 26 | 6  | 8  | 29 | 42 | -13 | 0.0%  | 0.0%  | FLA        | 38 | 9  | 11 | 36 | 39 | -3  | 0.0%  | 0.0%  | BOT        | 50 | 11 | 17 | 61  | 70 | -9  | 0.0%  | 0.0%  |
| GRE        | 15 | 3 | 6 | 21 | 25 | -4  | 0.4%  | 0.1%  | BOT        | 25 | 5  | 10 | 31 | 38 | -7  | 0.0%  | 0.0%  | PAR        | 37 | 10 | 7  | 39 | 61 | -22 | 0.0%  | 0.0%  | CAM        | 50 | 11 | 17 | 57  | 66 | -9  | 0.0%  | 0.0%  |
| PAY        | 13 | 3 | 4 | 20 | 27 | -7  | 0.2%  | 0.1%  | PAR        | 23 | 6  | 5  | 24 | 48 | -24 | 0.0%  | 0.0%  | VIT        | 36 | 10 | 6  | 55 | 67 | -12 | 0.0%  | 0.0%  | CRI        | 49 | 13 | 10 | 58  | 75 | -17 | 0.0%  | 0.0%  |
| BOT        | 13 | 2 | 7 | 20 | 27 | -7  | 0.2%  | 0.1%  | FLA        | 23 | 5  | 8  | 21 | 29 | -8  | 0.0%  | 0.0%  | SCA        | 35 | 17 | 8  | 45 | 30 | 15  | 0.0%  | 0.0%  | VIT        | 48 | 13 | 9  | 67  | 85 | -18 | 0.0%  | 0.0%  |
| FLA        | 12 | 2 | 6 | 16 | 23 | -7  | 0.1%  | 0.1%  | GUA        | 23 | 5  | 8  | 18 | 32 | -14 | 0.0%  | 0.0%  | GUA        | 46 | 10 | 16 | 41 | 55 | -14 | 0.0%  | 0.0%  | GUA        | 46 | 10 | 16 | 41  | 55 | -14 | 0.0%  | 0.0%  |
| SCA        | -4 | 5 | 5 | 18 | 14 | 4   | 0.0%  | 0.0%  | SCA        | 15 | 11 | 6  | 29 | 22 | 7   | 0.0%  | 0.0%  | GUA        | 31 | 6  | 13 | 29 | 46 | -17 | 0.0%  | 0.0%  | GRE        | 39 | 9  | 12 | 60  | 78 | -18 | 0.0%  | 0.0%  |

\*Professor do Depto. de Informática, FAMERC