

# Por que as zebras são comuns no Brasileirão

Modelo proposto por físicos mostra que aleatoriedade é alta em campeonatos de pontos corridos

Igor Zolnerkevic

Antes do início da edição atual do Campeonato Brasileiro, no final de maio, grande parte dos comentaristas esportivos apontava como favoritos ao título os times do Atlético Mineiro, do Corinthians e do Internacional, pela quantidade superior de bons jogadores que haviam contratado. O Cruzeiro também havia ganhado reforços, mas muitos não esperavam que no fim de setembro a raposa mineira fosse a líder do Brasileirão, seguida no *ranking* por Botafogo, Atlético Paranaense e Grêmio. A diferença de pontos entre os primeiros colocados, porém, é pequena. Tudo pode mudar até o final do torneio, em dezembro.

“Sabemos que nem sempre o melhor time ganha”, diz o físico Roberto da Silva, da Universidade Federal do Rio Grande do Sul (UFRGS). “No futebol há um grau de aleatoriedade muito alto.” Com a ajuda de colegas da UFRGS, Silva criou um modelo computacional que gera campeonatos virtuais com propriedades estatísticas idênticas àsquelas das pontuações de campeonatos nacio-

nais disputados por pontos corridos, como o brasileiro, o espanhol e o italiano. Seus resultados, publicados este ano na *Computer Physics Communications* e na *Physical Review E*, sugerem que as diferenças entre as habilidades dos times são importantes, mas que o que domina a dinâmica do futebol é mesmo a aleatoriedade, a tal caixinha de surpresas.

“Apostas em futebol não fazem ninguém rico”, diz Silva, que é natural de Mauá, na Grande São Paulo, e santista fanático. “É enorme o número de zebras, em que o pior ganha do melhor.” Ele justifica essa afirmação citando um estudo feito por estatísticos britânicos. O trabalho mostrou que os comentaristas esportivos de três jornais do Reino Unido acertaram só 42% das previsões feitas para 1.700 jogos do campeonato inglês. Quem apostasse sempre no time da casa, cuja chance de vitória costuma ser ligeiramente maior por conta da familiaridade com o gramado e a pressão da torcida, teria acertado 47%.

Silva e seus colegas analisaram as tabelas de pontuação de cinco edições do

Brasileirão e de vários campeonatos nacionais europeus disputados de 2006 a 2011. Desde 2003, o campeonato brasileiro adota o sistema de pontos corridos, como os europeus. Não há eliminatórias: cada um dos 20 times participantes joga duas vezes com os oponentes – uma vez em casa e a outra na casa do adversário. Os times ganham 3 pontos a cada vitória, 1 ponto por empate e nenhum na derrota. Cada time disputa um total de 38 partidas. É campeão quem somar o maior número de pontos ao final do torneio.

“Procurei na física um fenômeno aleatório que fosse similar”, Silva explica. Um dos processos aleatórios mais simples que existem é a difusão das moléculas de um soluto em um solvente, como acontece quando uma pitada de açúcar se dissolve num copo d’água. Silva tentou descrever a evolução das pontuações dos times com as mesmas equações que representam o movimento das moléculas na difusão. Nesse primeiro modelo, cada time seria uma molécula. O deslocamento de cada molécula corresponderia ao avanço das equipes ao longo do campeo-



Atlético Paranaense: em setembro, um dos líderes inesperados do campeonato brasileiro

nato, que podia se dar por três tipos de passo: derrota, empate ou vitória.

Não funcionou muito bem. Silva notou que a diferença entre as pontuações dos times tendia a aumentar mais rápido do que o projetado pelo modelo de difusão simples. Na verdade, a evolução dos pontos tinha as características daquilo que os físicos conhecem como superdifusão. Era o sinal de que a premissa do modelo simples, de que o desempenho dos times permanecia constante ao longo do campeonato, não se ajustava à realidade. “A superdifusão acontece quando as probabilidades de ganhar e perder se alteram ao longo do tempo”, explica. “Os times mudam: jogadores se machucam, novos jogadores são contratados e técnicos são demitidos.”

Num domingo, jogando *videogame* com o filho de 8 anos, Silva pensou em um modo de incorporar essas mudanças ao seu modelo. Como em um *videogame* de futebol, os times do modelo de Silva agora teriam um número que mediria a habilidade do grupo, ou seja, o seu potencial de ganhar uma partida. Os resultados

dos jogos continuavam a ser decididos aleatoriamente, mas a probabilidade de uma equipe ganhar ou perder passou a depender dos potenciais de ambos os times. Assim, o vencedor de uma partida aumentava seu potencial, enquanto o do perdedor diminuía. No empate, os potenciais dos times permaneciam constantes.

#### VIRTUAL X REAL

Com esse ajuste, o modelo funcionou melhor. Simulou com precisão a estatística acumulada de cinco campeonatos seguidos, mas apenas para o caso brasileiro. A pontuação dos torneios virtuais não batia com as dos campeonatos europeus, em especial do espanhol e do italiano.

Não foi difícil achar a explicação. O modelo assumia que todos os times começavam o campeonato com o mesmo potencial de vencer os jogos. Desde 2003, seis times já conquistaram o Brasileiro. Apesar de sempre haver favoritos, nenhum time nacional se destaca dos outros por muito tempo, por causa da habitual venda de craques para o exterior. Na Espanha é diferente. Os dois

melhores times – Barcelona e Real Madrid – têm uma média de gols por partida muito superior à das demais equipes e quase sempre Real Madrid ou Barça é campeão. O mesmo acontece na Itália, com Juventus, Milan e Internazionale. O modelo só funcionou para todos os países quando Silva incluiu essa diferença inicial, ajustando os potenciais dos times a partir da média de gols por partida de cada equipe no campeonato anterior.

Por ora, o modelo só reproduz de modo geral como a colocação dos times evoluiu. Mas Silva e seus colegas esperam conseguir rastrear um time individual e simular seu desempenho, avaliando suas chances de ganhar o campeonato.

O físico Haroldo Ribeiro, da Universidade Estadual de Maringá, também vem observando a superdifusão em suas análises de partidas de futebol, críquete e xadrez. “Ainda há muito o que investigar”, diz. “Podemos responder questões que os fãs dos esportes se perguntam ou justificar afirmações que muitas vezes eles fazem sem base científica.” ■

---

#### Artigos científicos

SILVA, R. *et al.* Anomalous diffusion in the evolution of soccer championship scores: Real data, mean-field analysis, and an agent-based model. **Physical Review E**. v. 88, n. 2. ago. 2013.

SILVA, R. *et al.* A simple non-Markovian computational model of the statistics of soccer leagues: Emergence and scaling effects. **Computer Physics Communications**. v. 184, n. 3. mar. 2013.