Medindo o tamanho da caixinha de surpresas em ligas de futebol

Raquel Y.S. Aoki, Renato M. Assunção, Pedro O.S. Vaz de Melo

[1]Departamento de Ciência da Computação – Universidade Federal de Minas Gerais (UFMG)
Belo Horizonte – MG – Brazil

{raquel.aoki,assuncao,olmo}@dcc.ufmg.br

Abstract. *Victory in soccer is a mixture of skill and luck. But how much of a victory is due to skill and luck? To answer this question, in this paper, we analyze soccer leagues that follow the double round-robin tournament system. We developed a scale with base in three reference points where we located a measure, also made in this paper, of observed variability in a championship. In total, we analyzed ten years of data from six championships. This analysis quantifies the notion that football is a box of surprises and allows one to measure to what extent a championship is more surprising than others.*

Resumo. *A vitória no futebol é uma mistura de habilidade e sorte. Mas quanto de sorte e de habilidade há em uma vitória? A fim de responder a esta pergunta, neste artigo analisamos campeonatos de futebol que seguem o sistema de pontos corridos. Desenvolvemos uma escala baseada em três pontos de referência onde localizamos uma medida, também produzida neste trabalho, da variabilidade observada em um campeonato. Ao todo, analisamos os dados de dez temporadas de seis campeonatos. Esta análise permite avaliar em que medida um campeonato é mais surpreendente que outro e quantifica a noção de que o futebol é uma caixinha de surpresas.*

1. Introdução

O futebol é um dos esportes mais populares do mundo e seus campeonatos são acompanhados por milhões de pessoas anualmente. A última Copa do Mundo, principal competição do esporte que é realizada a cada 4 anos, teve lugar no Brasil em 2014, com uma audiência estimada de 3.2 bilhões de pessoas [FIFA 2015]. Além dos campeonatos entre seleções, como a Copa do Mundo, todos os anos ocorrem campeonatos entre equipes. No Brasil, por exemplo, temos a Série A, na Inglaterra, a *Premier League*, na Espanha, a *Primera División*, nos Estados Unidos, a *Major League Soccer* e no Japão temos a *J-League*. O objetivo de todos esses campeonatos é o mesmo: eleger a melhor equipe . Mas será que a equipe campeã é realmente merecedora e mais habilidosa que as demais ou ela apenas teve sorte ao longo do campeonato? A resposta para essa pergunta não é óbvia, pois o resultado da maioria dos campeonatos é uma mistura de sorte e habilidade em quantidades desconhecidas.

Nosso trabalho foi motivado pelas análises desenvolvidas em [Spiegelhalter 2007], que propôs uma forma de medir a porcentagem de variação dos pontos de um campeonato que é devida à aleatoriedade. Entretanto, esta medida é pouco intuitiva. Ele é definida como uma porcentagem da variabilidade total, e deveria

estar sempre entre 0% e 100%. Entretanto, verifica-se empiricamente e teoricamente que esta medida pode facilmente extrapolar 100%. Quando isto acontece, o autor considera que o resultado do campeonato foi totalmente devido ao acaso. Mostramos que isto não é verdade, sendo esta uma grande desvantagem da medida proposta por [Spiegelhalter 2007].

Neste artigo, desenvolvemos uma escala para comparar a variabilidade observada da distribuição final dos pontos em um campeonato com a variabilidade teórica derivada de três modelos de referência. Um desses modelos determina o máximo da escala e descreve um campeonato com resultados completamente definidos pelas diferentes habilidades dos times, sem nenhum espaço para o acaso ou fatores externos tais como a vantagem como mandante. Num campeonato assim, os times nem precisam entrar em campo para saber o resultado do jogo. Um segundo modelo determina um ponto central e descreve um campeonato em que os times possuem as mesmas habilidades e os resultados dos jogos são determinados pelo puro acaso. Um terceiro modelo fixa o mínimo da escala e representa um campeonato em que os times terão exatamente a mesma pontuação, uma igualdade perfeita. Este caso extremo possui menos variabilidade do que o caso aleatório. Baseado nesta escala, obtemos uma medida que localiza a variabilidade observada em um campeonato em relação a estes pontos de referência.

Calculamos nossa medida em várias temporadas da Série A e Série B do campeonato brasileiro, em temporadas da *Major League Soccer* (MLS) dos Estados Unidos, na *División 1* da Argélia, na *Serie A* da Itália e da *Primeira División* da Espanha. Veremos que a maior parte dos resultados finais dos campeonatos são devido as diferenças de habilidade entre os times, exceto no caso do campeonato americano e argelino. De forma surpreendente, em alguns anos nestes campeonatos, os times estavam tão equilibrados entre si que dificilmente a habilidade de cada time ou o acaso tiveram um papel relevante.

2. Trabalhos Relacionados

Existem muitos trabalhos cujo objetivo é estudar campeonatos esportivos, em particular, os de pontos corridos. Em [Ben-Naim et al. 2013] é proposto um método para calcular a probabilidade do pior time vencer a competição. Os autores de [Chetrite et al. 2015] investigam o número de potenciais ganhadores nessas competições a partir de como são distribuídas as suas habilidades. Ainda neste tópico, [Ben-Naim et al. 2006] e [Goddard and Sloane 2014] estudam a competitividade de ligas esportivas. O primeiro avalia a competitividade a partir de uma medida de previsibilidade, concluindo que, quanto menos previsível for um torneio, maior é a competitividade entre as equipes participantes. O segundo estudo faz uso do conceito de *competitive balance* e explica como obter medidas para esse conceito.

Uma forma de estudar se o resultado de um campeonato é devido às habilidades das equipes ou ao acaso foi proposta em [Spiegelhalter 2007]. Ele compara a variância amostral observada e a variância teórica esperada no caso em que todas as equipes possuem a mesma habilidade, de onde se obtêm o percentual da variância da pontuação final do campeonato que é devida ao acaso. Essa metodologia só pode ser aplicada em campeonatos de pontos corridos, em que todas as equipes jogam entre si e o vencedor é a equipe que tiver a maior soma de pontos no final.

3. Metodologia

Devido à grande heterogeneidade que normalmente existe entre as habilidades das equipes de um campeonato, dificilmente será observado um torneio cujo vencedor seja o time de menor habilidade [Chetrite et al. 2015]. Entretanto, ocasionalmente pode ocorrer um equilíbrio entre as habilidades das equipes, seja porque todos os times são igualmente muito habilidosos, seja porque todos são pouco habilidosos. Nesse caso em que todas as equipes possuem as mesmas habilidades, o campeonato pode ser totalmente definido pelo acaso no sentido em que todas as equipes possuem a mesma probabilidade de vencer.

3.1. Probabilidade dos resultados

Para estimar a variância teórica do campeonato, [Spiegelhalter 2007] considera os N jogos ocorridos ao longo do campeonato e separa as W_m vitórias das equipes mandantes das W_v vitórias das equipes visitantes. Vamos estimar a probabilidade da equipe mandante vencer P_m por W_m/N e a probabilidade da equipe visitante vencer P_v por W_v/N. A probabilidade de empate vai ser estimada por subtração $P_e = 1 - P_m - P_v$.

Uma característica importante dos campeonatos de pontos corridos é que todas as equipes disputam o mesmo número de jogos como mandantes e visitantes. A Figura 1 mostra a probabilidade estimada dos resultados possíveis para cada temporada dos campeonatos estudados. Observando a Figura 1 é possível notar que uma equipe teria vantagens sobre as demais se disputasse mais jogos como mandante, pois o resultado mais provável de ser observado é a equipe mandante vencer. Além disso, a Figura 1 justifica porque estimar P_m, P_v e P_e separadamente.

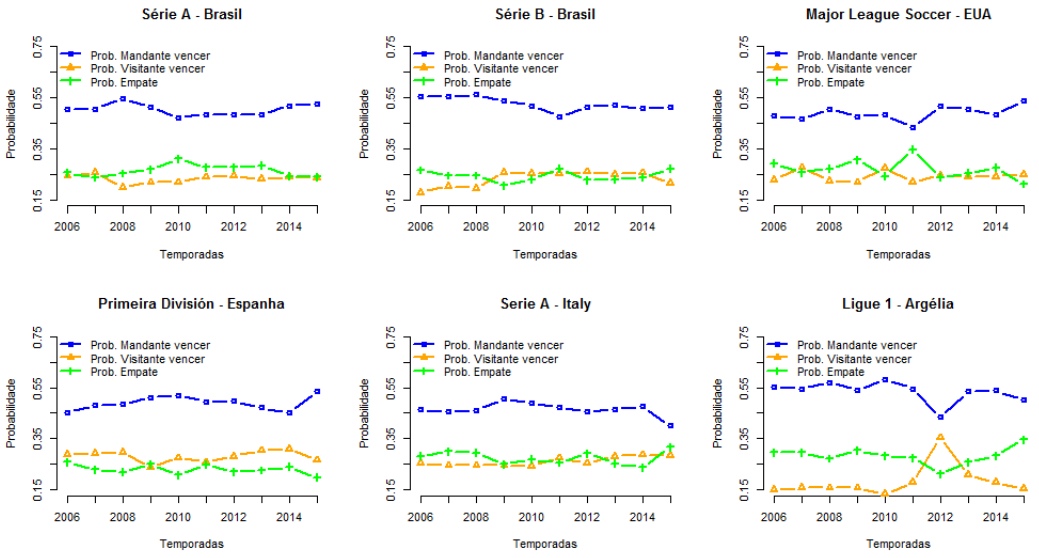

Figura 1. Probabilidade dos resultados possíveis de uma partida para os 6 torneios avaliados ao longo de 10 temporadas.

Vamos modelar os pontos ganhos por cada equipe na situação em que os times possuem habilidades idênticas. Considere duas variáveis aleatórias, uma que representa o número de pontos ganhos pela equipe como mandante e outra para os pontos ganhos como visitante. A variável aleatória X_m representa os pontos obtidos por um time em um

jogo como mandante, tendo média $\mu_m = 3 * P_m + P_e$ e variância $\sigma_m^2 = 9 * P_m + P_e - \mu_m^2$. Da mesma forma, X_v representa é o número de pontos ganhos como visitante e tem média $\mu_v = 3 * P_v + P_e$ e variância $\sigma_v^2 = 9 * P_v + P_e - \mu_v^2$.

Para unir as informações contidas em X_m e X_v foi criada uma outra variável aleatória, Y_{2k}, que representa o número de pontos ganhos por um determinado time após $2k$ partidas, k delas como mandante e as outras k como visitante. Assim, $Y_{2k} = X_{m1} + \ldots + X_{mk} + X_{v1} + \ldots + X_{vk}$ onde X_{mi} e X_{vi} são os pontos obtidos na i-ésima partida como mandante e como visitante, respectivamente. Note que, em campeonatos de pontos corridos, todas as equipes disputam o mesmo número de jogos como mandante e como visitante. Para uma amostra suficientemente grande de jogos e considerando que os jogos são independentes, podemos aproximar a soma Y_{2k} por uma distribuição gaussiana $N(\mu_{2k}, \sigma_{2k}^2)$ com $\mu_{2k} = k(\mu_m + \mu_v)$ e a variância $\sigma_{2k}^2 = k(\sigma_m^2 + \sigma_v^2)$. Essa é a distribuição teórica esperada da pontuação final das equipes no campeonato, após $2k$ jogos, caso todas possuam a mesma habilidade.

3.2. Percentual de variância devido ao acaso

Como proposto por [Spiegelhalter 2007], medimos a distância, em termos probabilísticos, entre a distribuição empírica dos pontos ganhos pelas equipes ao final de um campeonato e a distribuição teórica sob o modelo de igual habilidade dos times. Isto é obtido pela razão entre a variância teórica σ_{2k}^2 sob o modelo de habilidades iguais e a variância empírica s^2 dos pontos obtidos pelos times ao final do campeonato. Especificamente, se $T_1, T_2, ..., T_n$ é a pontuação dos n times ao final do campeonato, temos $s^2 = (T_1^2 + T_2^2 + ... + T_n^2)/n - \bar{x}^2$ onde $\bar{x} = (T_1 + T_2 + ... + T_n)/n$. Assim, define-se $\%Var = 100\% \times \sigma_{2k}^2/s^2$. Quando σ_{2k}^2 for maior que s^2, a medida $\%Var$ assumirá valores maiores que 100%. Esta situação pode ocorrer, e ocorre na prática, de fato ([Shergold 2015]), quando os times obtiverem pontuações muito similares entre si ao final do campeonato. As pontuações ficam mais parecidas entre si do que o modelo de aleatoriedade prediz através de σ_{2k}^2. Assim, interpretar esse percentual torna-se muito problemático.

3.3. Nossa proposta

A raiz do problema com a medida proposta em [Spiegelhalter 2007] é o uso da variância sob o modelo de aleatoriedade completa como se fosse um mínimo de variabilidade. Entretanto, ela não é nem um mínimo, nem um máximo. Para resolver este problema, calculamos a variância teórica máxima e mínima que um determinado campeonato pode ter e fizemos uma transformação linear para uma escala de -1 a 1. A medida σ_{2k}^2 será um valor intermediário e corresponde à situação em que os times não possuem habilidades diferentes.

A variância mínima é 0 e ocorre quando todas as equipes possuem exatamente o mesmo número de pontos. Esta é uma situação extrema, em que os resultados não são compatíveis com resultados casuais. No caso de um campeonato decidido pela pura sorte, deveríamos ter uma variância igual a σ_{2k}^2. Um valor muito menor pode aparecer como efeito de conluio entre os times ou algum outro mecanismo "compensatório"ao longo do campeonato. A variância máxima é mais complicada de ser calculada devido às particularidades de cada campeonato. Em um torneio em que todas as equipes jogam entre si e tem o mesmo número de jogos como mandante e visitantes, uma estimativa da variância máxima é obtida quando um campeonato é totalmente determinístico: os n

times são ordenados de acordo com suas habilidades e a primeira equipe ganha de todos os demais, a segunda equipe ganha de todos os demais com exceção da primeira equipe; e assim sucessivamente. Esse valor depende do número de times e quantidade de rodadas do campeonato e será chamado de MAX_D. Para colocar os valores de diferentes temporadas todos em uma mesma escala, foi feita a seguinte transformação linear:

$$\phi = \begin{cases} \frac{s^2 - \sigma^2_{2k}}{MAX_D - \sigma^2_{2k}}, & \text{se } \sigma^2_{2k} \leq s^2 \\ \frac{s^2 - \sigma^2_{2k}}{\sigma^2_{2k}}, & \text{caso contrário} \end{cases} \tag{1}$$

De maneira geral, valores positivos na escala indicam que o campeonato é definido mais pela habilidade das equipes que pela sorte, valores próximos de zero indicam que o torneio segue o modelo teórico em que as equipes possuem a mesma habilidade e valores negativos indicam que os pontos das equipes são mais similares entre si que o modelo aleatório pressupõe. Esta é uma situação curiosa: não há evidência de que as equipes sejam diferentes em suas habilidades e, ao mesmo tempo, o acaso não influencia muito os jogos. A exagerada similaridade entre as pontuações finais das equipes sugere que há, em tal campeonato, um mecanismo externo compensatório forçando uma contração em direção a uma pontuação igual para todos.

4. Bases de Dados e Resultados

A fim de avaliar a medida ϕ proposta neste trabalho, estudaremos os campeonatos Série A (Brasil), Série B (Brasil), *Major League Soccer* (Estados Unidos), *Serie A* (Itália), *Ligue 1* (Argélia) e a *Primeira División* (Espanha) ao longo de 10 temporadas. Nessa análise será estudada somente a primeira parte da *Major League Soccer* (MLS) que é de pontos corridos. Os dados utilizados nas análises foram coletados do site `http://www.betexplorer.com/soccer` no dia 29 de Junho de 2016.

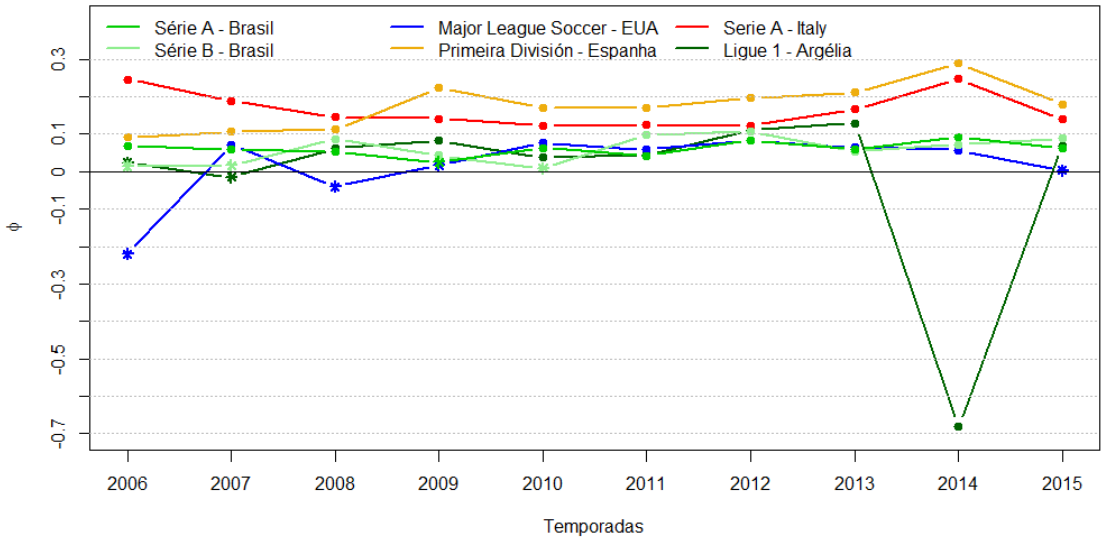

Figura 2. Escala da variância devido ao acaso dos campeonatos estudados.

A Figura 2 mostra o valor estimado de ϕ das 10 últimas temporadas dos campeonatos Série A, Série B, *Major League Soccer*, *Serie A*, *Ligue 1* e *Primeira División*. Para

verificar a significância estatística, simulamos os campeonatos sob o modelo de aleatoriedade completa dos resultados. Os valores de ϕ significativamente diferentes de zero ao nível de 5% são rotulados com $\bullet$ e os demais com $*$. Os torneios com os maiores valores de ϕ significativamente diferentes de zero foram a *Primeira División* e a *Serie A* italiana. Isso indica que o resultado final nesses campeonatos são primordialmente definidos pelas habilidades das equipes e não pelo acaso. Em relação aos campeonatos brasileiros, desde 2011 a Série A e Série B apresentam valores de ϕ entre 0.04 e 0.11 e significativamente diferentes de zero. A MLS apresentou quatro temporadas (2006, 2008, 2009 e 2015) com ϕ igual a zero, logo pode-se concluir que nesses anos as habilidades das equipes eram muito similares e o resultado do campeonato foi decidido totalmente devido acaso. A temporada da *Ligue 1* da Argélia em 2014 teve um ϕ significativo igual -0.68. A pontuação das equipes nesse ano era tão parecida que, à quatro rodadas do final, todas ainda tinham chances de vencer o campeonato [Shergold 2015]. Esse é um exemplo de torneio em que as equipes obtiveram uma pontuação mais similar do que a que o modelo aleatório prediz.

5. Conclusão

Neste artigo propomos uma medida ϕ para localizar a variabilidade numa escala onde os pontos extremos e central estão associados com modelos teóricos de campeonatos de pontos corridos. Ela permite avaliar a competitividade e o peso do acaso nos resultados observados. Dos seis campeonatos estudados, os menos surpreendentes são a *Primeira División* da Espanha e a *Serie A* da Itália. A Série A e a Série B do Brasil são mais surpreendentes que o campeonato espanhol e italiano, mas ainda assim, seus resultados se devem mais à habilidade das equipes que o acaso. A *Major League Soccer* apresentou quatro temporadas quase que totalmente definidas ao acaso e em 2014 a *Ligue 1* teve uma temporada em que as equipes tinham mais similaridades que o modelo aleatório supõe.

Referências

Ben-Naim, E., Hengartner, N., Redner, S., and Vazquez, F. (2013). Randomness in competitions. *Journal of Statistical Physics*, 151(3-4):458–474.

Ben-Naim, E., Vazquez, F., and Redner, S. (2006). Parity and predictability of competitions. *Journal of Quantitative Analysis in Sports*, 2(4).

Chetrite, R., Diel, R., and Lerasle, M. (2015). The number of potential winners in bradley-terry model in random environment. *arXiv preprint arXiv:1509.07265*.

FIFA (2015). 2014 fifa world cup^{TM} reached 3.2 billion viewers, one billion watched final. [Online; accessed 24-June-2016].

Goddard, J. and Sloane, P. (2014). *Handbook on the economics of professional football*. Edward Elgar Publishing.

Shergold, A. (2015). Algerian league is so tight all 16 teams can mathematically still win the title with four rounds of matches to go.

Spiegelhalter, D. (2007). Football leagues. [http://understandinguncertainty.org/node/314; accessed 26-June-2016].