

É o futebol um processo de difusão?



Aluno: Felipe S. F. Paula¹
Orientador: Roberto da Silva
Co-orientador: Sebastián Gonçalves

Resumo

O futebol tem um grau de aleatoriedade grande. Haja visto que comentaristas ingleses acertam apenas 42% dos resultados contra 47% de acertos quando simplesmente se aposta no time casa. Neste trabalho, temos como objetivo, modelar os aspectos difusivos deste esporte praticado em mais de 200 confederações, através de um simples modelo multi-agente. Neste modelo alimentando um algoritmo com o número médio de gols de cada time de um campeonato anterior nós obtivemos as pontuações de 5 campeonatos futuros e com isso pudemos calcular as propriedades estatísticas das flutuações emergentes do nosso campeonato simulado. Também obtivemos um caso particular deste modelo o que corresponde a uma aproximação de campo médio dele. Nossos resultados foram comparados com os obtidos de 5 edições de diferentes campeonatos no mundo. Comprovamos uma excelente concordância entre as distribuições de pontuações de dados reais com os dados simulados mostrando que o futebol apresenta comportamento superdifusivo tanto pelos dados reais quanto pelo nosso modelo multi-agente.

Modelo baseado em agentes

Seja $\varphi_t^{(i)}$ o potencial do i -ésimo time no instante t e $r_{draw}^{i,j}$ a probabilidade de empate. O time i ganha do time j com probabilidade

$$\Pr(i > j, t) = [1 - r_{draw}^{(i,j)}(t)] \frac{\varphi_t^{(i)}}{(\varphi_t^{(i)} + \varphi_t^{(j)})}$$

Sabemos que o número de gols marcados por um time durante uma partida, pode ser modelada através da distribuição de Poisson, onde a taxa de ocorrências representa o potencial do time. Então a probabilidade de dois times terem o mesmo score, isto é, empate, pode ser dada por:

$$r_{draw} = \Pr[(n_i = n_j) | (\varphi_i^{(i)}, \varphi_i^{(j)})] = \sum_{n=0}^{\infty} \frac{\varphi_i^{(i)n}}{n!} e^{-\varphi_i^{(i)}} \frac{\varphi_i^{(j)n}}{n!} e^{-\varphi_i^{(j)}} \\ = e^{-(\varphi_i^{(i)} + \varphi_i^{(j)})} I_0(2\sqrt{\varphi_i^{(i)} \varphi_i^{(j)}})$$

Onde $I_\nu(z)$ é a função de Bessel de primeiro tipo modificada. Quando um time ganha, o seu potencial é aumentado pelo seu mesmo valor dividido pelo número de rodados, isto é, $\varphi \rightarrow \varphi + \varphi/k$, caso perca é decrementado da mesma quantidade e em caso de empate continua o mesmo.

Coeficientes obtidos

País	C_{real}	C_{alg}	C_{cm}	β_{real}	β_{alg}	β_{cm}	D_{real}	D_{alg}	D_{cm}
Brasil	1.37	1.375	1.364	0.60	0.593	0.544	1.28	1.49	1.70
Espanha	1.38	1.364	1.363	0.72	0.680	0.545	1.18	1.10	1.70
França	1.35	1.358	1.340	0.66	0.583	0.542	1.25	1.46	1.60
Inglaterra	1.36	1.364	1.363	0.74	0.636	0.543	1.02	1.22	1.70

Conclusões

Podemos ver na tabela acima que ambas abordagens reproduzem muito bem o coeficiente C . O algoritmo (baseado em agentes), descreve muito bem o comportamento superdifusivo ($\beta > 1/2$) dos dados, contudo a aproximação de campo médio reproduz apenas parcialmente. Finalmente, podemos dizer que, baseando-se nos momentos e coeficientes obtidos, que o modelo proposto reproduz as flutuações do futebol.

Aproximação de campo médio

Vamos considerar aqui a probabilidade de empate $r_{draw}^{i,j}(t) = r$ e potencial dos times $\varphi_t^i = \varphi$ sem depender do tempo. Com isso, as probabilidades do times vencer $p_w = \Pr(i > j, t)$, perder $p_l = \Pr(i < j, t)$ e empatar se tornam constantes, isto é, $p_w = p_l = (1 - r)/2$. Denotando como $P_m(n)$ a probabilidade de se ter n pontos no instante m , podemos ter:

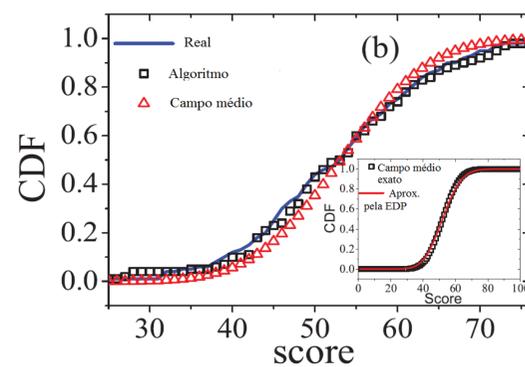
$$P_{m+1}(n) = r P_m(n - 1) + p_w P_m(n - 3) + p_l P_m(n).$$

Essa relação pode ser expandida se pegarmos $p_w = p_l = (1 - r)/2$. Depois disso usando os operadores de diferenças finitas e tomando o limite chegamos na seguinte equação de difusão.

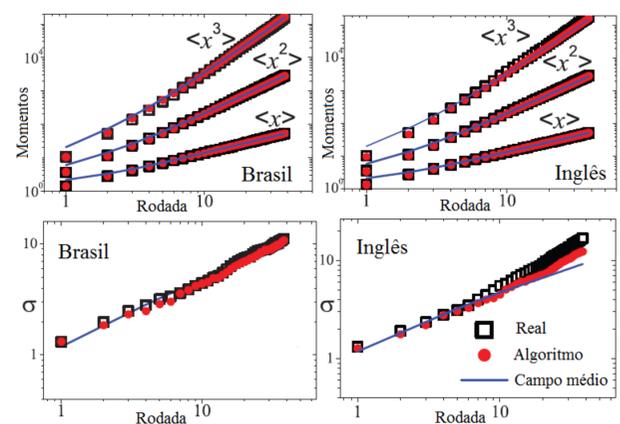
$$\frac{\partial P}{\partial t} = -C \frac{\partial P}{\partial x} + D \frac{\partial^2 P}{\partial x^2} - E \frac{\partial^3 P}{\partial x^3}$$

Onde $C = r + 3(1 - r)/2$, $D = 3(1 - r)/2$ e $E = (1 - r)/2$

Resultados



O gráfico acima mostra a função distribuição acumulada (CDF) dos dados reais (38ª rodada do Brasileirão) junto do previsto pelo modelo baseado em agentes (algoritmo) e campo médio. O gráfico mais interior mostra também a comparação do campo médio exato e o aproximado pela EDP.



No gráfico acima temos as flutuações reais e as obtidas através do algoritmo e aproximação de campo médio. Podemos ver que o modelo multi-agente se aproxima mais dos dados reais.

Referências

- R. da Silva, M. H. Vainstein, L. C. Lamb, S. D. Prado, A simple non-Markovian computational model of the statistics of soccer leagues: Emergence and scaling effects. *Computer Physics Communications*, 184, 3, 661670(2013)
- R. da Silva, M. H. Vainstein, S. Gonçalves, F. S. F. Paula, Anomalous diffusion in the evolution of soccer championship scores: Real data, mean-field analysis, and an agent-based model. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, v. 88, p. 022136-022136-8, 2013.