



A Matemática pelo Basquetebol

Será possível prever o resultado de uma partida de basquetebol? Além das análises dos dados estatísticos, como pode a matemática contribuir para o basquetebol? Curiosamente, a ferramenta que ordena as ligações apresentadas pelo motor de busca *Google* recorre a conceitos que podemos aplicar na previsão de um resultado para uma partida de basquetebol.

1. Introdução

A ideia de recorrer a computadores para cálculos matemáticos aplicados ao basquetebol já existia, pelo menos, desde 1958, quando Donald Knuth se envolveu com a equipa de basquetebol *Case* na aplicação de uma fórmula para atribuir valores aos jogadores. Com o impulso dado por este caso, e por se pretender para o basquetebol português mais matemática do que a análise das estatísticas de jogo, foi-se concretizando a ideia da aplicação da matemática a esta modalidade, que acabou por tomar forma numa tese de mestrado¹. Este artigo pretende ter um carácter de divulgação matemática, mostrando como a matemática se pode aplicar a um desporto, com algo mais do que o cálculo de percentagens e médias.

2. O basquetebol

Apesar de existirem referências à ideia de um jogo/ritual, cujo objectivo era introduzir um objecto dentro de um aro, nas civilizações Maia e Azteca, o basquetebol como desporto é tido como tendo surgido em 1891.

Em Dezembro de 1891, o presidente do Departamento de Educação Física da School for Christian Workers² incumbiu o professor de Educação Física James Naismith de inventar um jogo para entreter os atletas da escola durante a época de Inverno.

¹Ver [2].

²Actualmente Springfield College.

³Os cinco princípios iniciais que iriam definir a nova actividade podem ser vistos em [3].

⁴As bolas de futebol só foram substituídas por bolas de basquetebol em 1894.

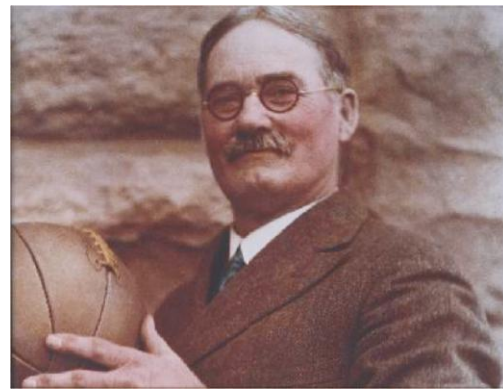


Figura 1: James Naismith

Depois de reflectir sobre os princípios fundamentais da sua nova actividade³, decidiu testar o seu jogo. A caminho do ginásio, Naismith pediu ao porteiro duas caixas, que serviriam de alvo para a bola (que começou por ser uma bola de futebol)⁴. O porteiro respondeu: “Tenho dois cestos de apanhar pêssegos! Se servirem de alguma coisa...” Apresentava-se assim o jogo constituído por dois cestos (*basket*) e uma bola (*ball*) – o *basketball*.

3. O que se estuda

O basquetebol é, actualmente, uma modalidade desportiva de nível mundial. Em [1] refere-se que este jogo “é praticado por mais de 11% da população

mundial, chegando inclusivamente a ultrapassar o número de praticantes de futebol em muitas regiões do globo”.

Dean Oliver³ tem vários artigos publicados no seu *site*⁶ onde apresenta diversas abordagens e análises ao jogo. Por exemplo, num artigo de Dezembro de 1995 – *Basketball's Bell Curve* – apresenta a distribuição dos pontos marcados por uma equipa. Mais especificamente, o autor refere-se à curva de Gauss como representação da distribuição dos pontos marcados pela equipa para jogos como visitante e jogos como visitado. Comparando as duas distribuições, poder-se-á inferir a vantagem do factor casa. As análises apresentadas permitem desenhar estratégias de jogo relevantes. Por exemplo, tendo em conta as observações feitas, uma equipa não favorita deverá apostar na variação ou inconsistência contra uma equipa favorita. Se for previsível para a equipa não favorita um resultado de 104 contra 106 da favorita, uma estratégia ofensiva a considerar será a aposta nos lances tripos, de forma a tentar beneficiar com o aumento do desvio padrão.



Figura 2: Cestos de pêssegos e bola

Uma abordagem interdisciplinar interessante é feita por Yilmaz e Chatterjee⁷. No seu estudo faz-se uma análise ao desempenho na NBA ao longo de cinco décadas. O ponto de vista considerado é que torna esta abordagem curiosa – o ponto de vista evolucionista da biologia. A teoria evolucionista defende que as espécies progridem,

através de mudanças ao longo do tempo, de organismos primitivos para seres mais complexos. O processo que envolve tais mudanças e que permite a sobrevivência dos mais aptos, eliminando os menos aptos, é designado por selecção natural. O biólogo Stephen Jay Gould recorreu a estas ideias para interpretar o desempenho no *baseball*. Tendo em conta os longos anos de evolução de tal desporto, é possível concluir que, tal como num sistema biológico, há uma moderação dos extremos à medida que o sistema adquire estabilidade. Os autores do artigo anteriormente referido examinaram, nesta perspectiva, o desempenho de equipa na NBA. Após a análise dos dados, Yilmaz e Chatterjee concluíram que, enquanto no *baseball* a variabilidade nas medidas de desempenho das equipas apresenta uma diminuição estável, na NBA tal estabilidade ainda não é observável. Tal como já mencionado, o tempo de vida maior do *baseball* profissional em relação ao do basquetebol e as poucas mudanças nas regras do primeiro desporto serão algumas das justificações para tais resultados. No entanto, como referem os mesmos autores em [4], o padrão verificado no *baseball* começa a emergir no basquetebol.

O basquetebol, sendo um desporto que envolve o lançamento de um objecto, está também sujeito a uma análise do ponto de vista da física. Uma ideia, talvez bastante comum, será a de que os jogadores profissionais de basquetebol, as grandes estrelas, serão, pela prática que têm, exímios marcadores nos lances livres. Gablonsky e Lang recorrem ao exemplo do conhecido jogador Shaquille O'Neal que, na época regular 2004/2005 do campeonato norte-americano, tinha uma eficácia de 53,1%. Tendo em conta os dois factores essenciais num lançamento livre – ângulo e velocidade – Gablonsky e Lang⁸ apresentam um modelo matemático para este tipo de lançamento. Baseados na segunda lei de Newton e nas equações de movimento de um projectil, os autores desenvolvem modelos, partindo de um modelo simples, bastante geral, até um modelo mais realista com mais condições. Uma conclusão que surge do estudo das margens de erro nos dois factores importantes no lançamento é a de que é mais importante ser preciso na velocidade do que no ângulo. Uma conclusão

³Actualmente, *director of quantitative analysis* na equipa americana Denver Nuggets.

⁶Ver [5].

⁷Cf. [6].

⁸Cf. [7].

⁹Imagem em <http://www-history.mcs.st-andrews.ac.uk/PictDisplay/Markov.html>

curiosa sobre a trajectória da bola, apresentada pelos autores, é a de que o jogador não deve, ao preparar o lançamento livre, apontar para o centro do cesto, mas sim para algum ponto entre o centro e a parte de trás do aro. Além disso, quanto mais baixo for o jogador, mais atrás deve ser esse ponto.

4. Cadeias de Markov

Em 1907 Andrei Andreyevich Markov (1856–1922)⁹ iniciou o estudo de processos aleatórios em que apenas o resultado da experiência mais recente afecta o resultado da experiência seguinte.

As cadeias de Markov, a partir daí desenvolvidas, são amplamente usadas em diversas áreas, inclusive no algoritmo *PageRank*, usado pelo motor de busca *Google* e desenvolvido pelos seus fundadores enquanto estudantes universitários¹⁰.

Definição

Considere-se um conjunto de estados $E = \{e_1, e_2, \dots, e_n\}$. Cada passo do processo consiste na passagem ou transição de um estado para outro. Estando a cadeia de estados actualmente no estado e_i , esta passa, no passo seguinte, para o estado e_j com uma probabilidade P_{ij} . À probabilidade P_{ij} chama-se *probabilidade de transição* do estado e_i para o estado e_j . A *propriedade de Markov* estabelece que o estado da cadeia num instante futuro, dado o conhecimento do estado da cadeia no passado e no presente, depende apenas do presente e não do passado. Simbolicamente, sendo a cadeia um conjunto de variáveis aleatórias X_1, X_2, X_3, \dots que tomam como valores os estados e_i , a propriedade descrita consiste no facto de que

$$\forall n \in \mathbb{N}, e_{i_1}, \dots, e_{i_n}, e_j \in E$$

$$P(X_{n+1} = e_j \mid X_n = e_{i_n}, X_{n-1} = e_{i_{n-1}}, \dots, X_1 = e_{i_1}) = P(X_{n+1} = e_j \mid X_n = e_{i_n})$$

Por outras palavras, a probabilidade de a cadeia se encontrar num determinado estado depende apenas do estado anterior. Em geral, começa-se por especificar um estado inicial para o processo ou um vector de probabilidades para tal estado.

¹⁰Um relatório técnico onde o *PageRank* é apresentado – *The PageRank Citation Ranking: Bringing Order to the Web* – pode ser encontrado em [8].

¹¹Recorde-se que teremos uma cadeia de Markov, pois admitimos que o estado do tempo de um determinado dia depende apenas do estado do tempo do dia anterior.

Exemplo

Imaginemos um local onde nunca há dois dias seguidos de bom tempo. Se num dia estiver bom (B) tempo, no dia seguinte há a mesma probabilidade ($\frac{1}{2}$) de haver chuva (C) ou neve (N). Havendo chuva ou neve, há uma probabilidade de $\frac{1}{2}$ de se repetir o tempo no dia seguinte. Por outro lado, estando chuva ou neve, se o tempo mudar no dia seguinte, só metade das vezes muda para bom tempo.

Com os dados anteriores pode-se construir um modelo de Markov¹¹. Sejam B, C e N os três estados de tempo possíveis. Com os dados fornecidos, facilmente se determinam as probabilidades de transição entre os estados. Com estas pode-se produzir uma matriz $P = [P_{ij}]$, em que P_{ij} representa a probabilidade de transição do estado i para o estado j :

$$P = \begin{matrix} & \begin{matrix} C & B & N \end{matrix} \\ \begin{matrix} C \\ B \\ N \end{matrix} & \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{pmatrix} \end{matrix}$$

Assim, por exemplo, a probabilidade de no dia seguinte a um dia de chuva (C) nevar (N) é igual a $P_{33} = \frac{1}{4}$.

A uma matriz como esta, que contém as probabilidades de transição, chama-se *matriz de transição*.

5. Um jogo de basquetebol

Para definir o modelo de Markov, considerou-se que a jogada de cada equipa pode percorrer seis possíveis estados, tendo-se assim, no total, os seguintes doze estados:

- e_1 : equipa A com posse de bola, zona de 3 pontos;
- e_2 : equipa A lança ao cesto, zona de 3 pontos;
- e_3 : equipa A marca;
- e_4 : equipa A falha;
- e_5 : equipa A com posse de bola, zona de 2 pontos;
- e_6 : equipa A lança ao cesto, zona de 2 pontos;
- e_7 : equipa B com posse de bola, zona de 3 pontos;
- e_8 : equipa B lança ao cesto, zona de 3 pontos;
- e_9 : equipa B marca;
- e_{10} : equipa B falha;

- e_{11} : equipa B com posse de bola, zona de 2 pontos;
- e_{12} : equipa B lança ao cesto, zona de 2 pontos.

O modelo de Markov, com os estados descritos anteriormente, pode ser representado pelo grafo apresentado na figura 3.

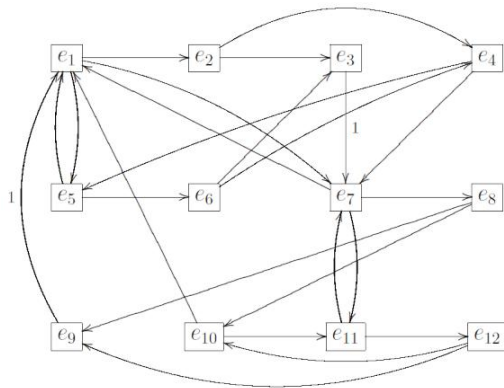


Figura 3: Transições possíveis

O modelo de Markov construído recorre a diversas probabilidades de transição. Por exemplo: P_{12} representa a probabilidade de a equipa A lançar ao cesto da zona de 3 pontos; P_{71} representa a probabilidade de a equipa B perder a bola (*turnover*) na zona de 3 pontos. Os valores utilizados para as probabilidades de transição foram estimados com base nas frequências relativas correspondentes.

6. O simulador

O simulador foi desenvolvido com base no software *Mathematica* e tem uma utilização bastante acessível, já que se desenvolveu para o efeito um interface gráfico para o utilizador. O seu aspecto, ao ser iniciada a simulação, é apresentado na figura 4.

Depois de introduzidas as estimativas das probabilidades de transição, o utilizador pressiona o botão "Jogar". O estado inicial da cadeia de Markov é um dos dois estados possíveis, e_i ou e_j , decidido de forma aleatória e com a mesma probabilidade para ambos. No final da simulação, uma mensagem indica que aquela terminou e é apresentado o resultado previsto para as duas equipas.



Figura 4: O simulador.

7. Simulação de jogos

Recorrendo à análise de dados disponíveis sobre as equipas de basquetebol portuguesas, estudaram-se alguns jogos e recorreu-se ao simulador para efectuar previsões antes dos jogos.

Um dos jogos para o qual se aplicou o simulador foi o Porto/Ovarense, ocorrido a 13 de Fevereiro de 2007, no âmbito da 24.ª jornada da época regular 2006/2007.

Estimadas as probabilidades de transição e definido o número de jogadas, tais elementos foram introduzidos no simulador. Uma das previsões produzidas¹² foi:

Porto	69,81
Ovarense	70,64

No final do jogo, verificou-se que o resultado (sem lances-livres) foi Porto-58, Ovarense-68. Verifica-se que a previsão relativa à equipa do Porto é superior à pontuação conseguida. No que diz respeito à equipa da Ovarense, a previsão é bastante próxima do verificado.

As previsões e os resultados reais eram, frequentemente, discutidos no final dos jogos com o treinador da equipa do Porto, o que permitia compreender as eventuais diferenças entre as duas pontuações e acentuava o carácter de ligação à prática deste projecto.

¹²Para cada jogo efectua-se previsões em diversos cenários, dependentes do número de jogadas total.

8. A partir daqui...

Este artigo ambiciona apenas abrir possíveis caminhos que podem agora ser prolongados pelo leitor que os pretenda aprofundar.

A compreensão do lance livre e o seu estudo físico-matemático, adaptado às equipas portuguesas, revelam-se como um campo de aplicação interessante.

Vários momentos do jogo de basquetebol são baseados em estratégias de acção. A teoria de jogos terá também, certamente, uma contribuição relevante a dar.

A geometria poderá contribuir para análises das equipas e da relação entre os jogadores. Enriquecerá as observações, já que difere bastante das análises convencionais feitas aos jogos. Nesta área, assim como noutras, a tecnologia poderá ser bastante útil. Tais ferramentas poderiam até auxiliar no aumento

de dados disponíveis como, por exemplo, o tempo de jogada ou o número de passes de bola por jogada.

O estudo das estatísticas das equipas e a comparação dos valores MVP (*most valuable player*) com as classificações das equipas constitui um passo a desenvolver. Aliás, até Donald Knuth sugeriu, em resposta a umas questões sobre os resultados obtidos "*Can you go back to ten or more previous seasons, to get more numbers?*"

Relativamente ao simulador desenvolvido, este constitui um ponto de partida que poderá ser desenvolvido para fornecer pontuações mais precisas ou outras informações. O próprio simulador seria mais completo (e certamente mais complexo) se tivesse em conta a influência de certos jogadores no desempenho de outros da mesma equipa, algo que passará despercebido numa leitura simples de muitos dados de um jogo. [\[1\]](#)

Agradecimento

O autor agradece as sugestões apresentadas pelo *referee*, por terem contribuído bastante para melhorar o presente texto.

Referências

- [1] **Sampaio, J.** (2000). "O poder discriminatório das estatísticas do jogo de basquetebol: novos caminhos metodológicos de análise.". Tese de Doutoramento, Departamento de Desporto da Universidade de Trás-os-Montes e Alto Douro.
- [2] **Machado, R.** (2007). "Basquetemática: Uma incursão da matemática pelo basquetebol.". Tese de Mestrado, Faculdade de Ciências da Universidade do Porto.
- [3] "Dr. James Naismith Foundation.". <http://epe.lac-bac.gc.ca/100/205/301/ic/cdc/naismith/index.htm>
- [4] **Chatterjee, S. & Yilmaz, M.** (1999). "The NBA as an evolving multivariate system.". *The American Statistician*, (3) 53.
- [5] **Oliver, D.** "Journal of Basketball Studies.". www.rawbw.com/~deano
- [6] **Yilmaz, M. & Chatterjee, S.** "Patterns of NBA team performance from 1950 to 1998." www.metapress.com/content/cc65d0x4wbxwllbc/?p=0e826fd40c1a4d13bd9510a4a3a21dee\&pi=2
- [7] **Gablonsky, J. & Lang, A.** (2005). "Modeling basketball free throws.". *SIAM Review*, (4) 47.
- [8] "Publication Server of the Stanford InfoLab." <http://ilpubs.stanford.edu:8090>