

Previsões para os jogos da Copa do Mundo de 2014

Paulo Cezar Pinto Carvalho, Moacyr Alvim Horta Barbosa da Silva, Arthur da Silva Pereira Carneiro
FGV/EMAp

Resumo: Um modelo probabilístico foi desenvolvido para fazer previsões sobre os jogos da Copa do Mundo de 2014. Para tal, admite-se que o número de gols que uma equipe A faz sobre uma equipe B tem uma distribuição de Poisson, em que a média depende de parâmetros de força de defesa e ataque, a serem estimados com base em todos os jogos entre seleções realizados nos últimos 4 anos. O modelo permite prever o resultado mais provável de cada partida e, via simulação, as probabilidades de sucesso de cada seleção.

O Modelo

Diversos autores sugeriram que o número de gols assinalados em uma partida de futebol pode ser modelado por uma distribuição de Poisson. Neste trabalho, adotamos uma modificação do método de Mahler (1982), modelando o número de gols marcados por uma equipe A em uma equipe B como tendo distribuição de Poisson com média $\frac{\alpha_A}{\beta_B}$, onde α_A é um parâmetro que mede a força de ataque da equipe A e β_B é um parâmetro que mede a força de defesa da equipe B. Os valores de α e β para as 211 seleções afiliadas à FIFA foram estimados por máxima verossimilhança, usando os resultados de todos os 3598 jogos entre seleções nos últimos 4 anos. Cada partida é afetada de um peso, que é o produto de dois fatores: o primeiro expressa a importância de cada partida e o segundo descreve quão recente ela é. Os pesos adotados para a importância são os mesmos usados pela FIFA na construção de seu ranking: 1 para amistosos, 2,5 para jogos classificatórios para a Copa do Mundo ou Copa Continentais, 3 para jogos de finais Continentais e da Copa das Confederações e 4 para a Copa do Mundo. Para o peso temporal, adotamos um amortecimento exponencial tal que o fator associado a um jogo realizado há 4 anos atrás é 0,25.

Os valores estimados de α e β permitem construir um ranking das equipes, associando a cada uma o valor $\lambda = \alpha \beta$, ou seja, o produto das forças de ataque e defesa. Este produto produz um ranking apropriado, no sentido de que se $\lambda_A > \lambda_B$, então A tem maior probabilidade do que B de vencer uma partida entre eles.

A tabela abaixo fornece as forças de ataque e defesa para as 32 seleções classificadas para as finais da Copa do Mundo de 2014, além de sua força combinada λ .

País	Força de ataque	Força de defesa	Força combinada
Brasil	3.9138	2.9501	11.5462
Croácia	1.5930	1.7091	2.7226
México	2.2926	2.0086	4.6049
Camarões	1.4953	1.2185	1.8220
Espanha	3.4269	2.3259	7.9706
Holanda	2.6371	1.9562	5.1586
Chile	3.0086	1.7979	5.4092
Austrália	2.2093	1.3547	2.9929
Colômbia	2.4679	3.2310	7.9738
Grécia	1.2988	1.8139	2.3559
Costa do Marfim	2.5480	1.3196	3.3624
Japão	2.4753	1.4531	3.5969
Uruguai	2.9749	1.7505	5.2075
Costa Rica	1.6217	1.9655	3.1875
Inglaterra	2.1740	1.9136	4.1601
Itália	2.5808	1.3264	3.4230
Suíça	2.2150	1.2531	2.7756
Equador	2.3914	1.9969	4.7755
França	2.1435	2.3331	5.0011
Honduras	1.7305	1.5320	2.6512
Argentina	3.4037	2.1036	7.1599
Bósnia	2.2294	1.3668	3.0472
Irã	2.0683	2.1272	4.3996
Nigéria	1.1534	1.0140	1.1695
Alemanha	3.6014	1.7116	6.1643
Portugal	2.3971	1.7143	4.1093
Gana	1.9456	1.5374	2.9912
Estados Unidos	2.6114	1.7555	4.5842
Bélgica	2.2967	1.8904	4.3417
Argélia	1.5681	1.5835	2.4830
Rússia	1.8793	2.4689	4.6399
Coréia do Sul	2.0421	1.5534	3.1722

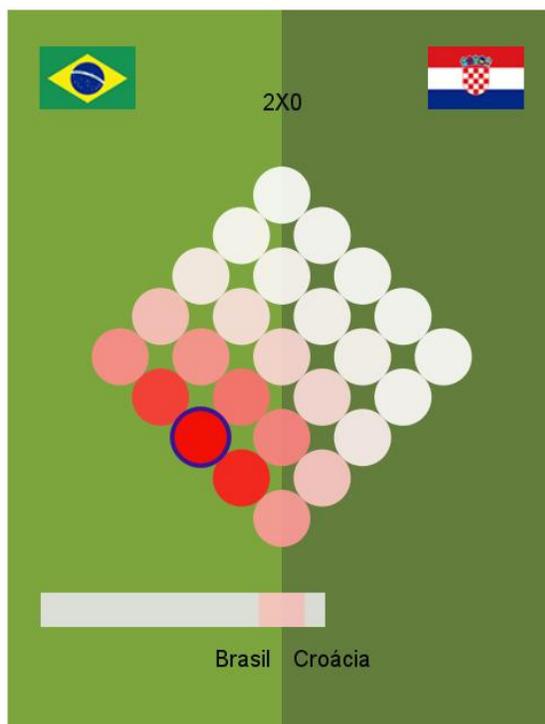
As previsões

Uma vez estimadas as forças de defesa e ataque, produzimos dois tipos de previsão:

- Para os jogos da primeira fase, calculamos as probabilidades associadas aos diversos resultados. Por exemplo, para o jogo de estreia do Brasil, contra a Croácia, os valores das probabilidades dos diversos resultados são dados na tabela abaixo.

		Brasil				
		0	1	2	3	4
Croácia	0	0.0217	0.0644	0.0955	0.0944	0.0700
	1	0.0188	0.0556	0.0825	0.0816	0.0605
	2	0.0081	0.0240	0.0356	0.0352	0.0261
	3	0.0023	0.0069	0.0103	0.0101	0.0075
	4	0.0005	0.0015	0.0022	0.0022	0.0016

Este mapa de probabilidades pode ser melhor visualizado graficamente. Optamos pela representação abaixo, em que a intensidade da cor é proporcional à probabilidade do resultado correspondente. A barra horizontal registra as probabilidades de vitória de cada seleção e de empate.



O resultado mais provável é Brasil 2 x 0 Croácia, com probabilidade aproximadamente igual a 9,6%. O resultado mais provável na partida A x B é simples de se obter e é dado por $\left\lfloor \frac{\alpha_A}{\beta_B} \right\rfloor \times \left\lfloor \frac{\alpha_B}{\beta_A} \right\rfloor$, onde $\lfloor x \rfloor$ denota o maior inteiro menor ou igual a x .

Os resultados de todas as previsões podem ser encontrados em emap.fgv.br/copa-2014.

- Usando as forças estimadas podemos simular um grande número de realizações da Copa do Mundo e obter as probabilidades de cada equipe ser campeã. Podemos também estimar as chances de uma equipe atingir cada fase do torneio. Os resultados são dados na tabela abaixo.

País	Probabilidade de ser campeão	Probabilidade de atingir...			
		as oitavas	as quartas	as semi-finais	a final
Brasil	29%	98%	74%	52%	40%
Croácia	0%	45%	9%	3%	1%
México	1%	44%	15%	8%	3%
Camarões	0%	13%	1%	0%	0%
Espanha	16%	98%	70%	48%	31%
Holanda	2%	45%	14%	7%	4%
Chile	2%	44%	14%	7%	4%
Austrália	0%	13%	2%	1%	0%
Colômbia	12%	98%	69%	32%	21%
Grécia	0%	45%	13%	3%	1%
Costa do Marfim	0%	44%	16%	4%	1%
Japão	0%	13%	5%	1%	1%
Uruguai	5%	98%	63%	26%	12%
Costa Rica	0%	45%	15%	3%	1%
Inglaterra	1%	44%	16%	5%	2%
Itália	0%	13%	4%	1%	0%
Suíça	1%	98%	41%	13%	3%
Equador	1%	45%	18%	10%	4%
França	2%	44%	19%	10%	4%
Honduras	0%	13%	3%	1%	0%
Argentina	12%	98%	65%	45%	24%
Bósnia	0%	45%	22%	7%	1%
Irã	1%	44%	26%	12%	4%
Nigéria	0%	13%	6%	2%	0%
Alemanha	9%	98%	68%	46%	18%
Portugal	1%	45%	22%	9%	3%
Gana	0%	44%	17%	5%	1%
Estados Unidos	0%	13%	7%	3%	1%
Bélgica	3%	98%	52%	21%	8%
Argélia	0%	45%	11%	4%	1%
Rússia	1%	44%	19%	11%	4%
Coréia do Sul	0%	13%	4%	2%	0%

Atualização das previsões

Todas as estimativas acima dependem das estimativas das forças de cada seleção. Tais valores mudam, à medida que novos amistosos e, principalmente, os jogos iniciais da Copa sejam realizados. Todos os resultados serão atualizados diariamente e publicados em emap.fgv.br/copa-2014.

Referência:

Maher M.J. (1982), *Modelling Association Football scores*. *Statistica Neerlandica*, 36, 109-118.