

UMA ABORDAGEM BAYESIANA PARA PREVISÃO DE RESULTADOS DE JOGOS DE FUTEBOL: UMA APLICAÇÃO AO CAMPEONATO INGLÊS

Ciro Alexandre Olivieri FILHO¹
Adriano Kamimura SUZUKI¹
Francisco LOUZADA¹
Erlandson Ferreira SARAIVA²
Luis Ernesto Bueno SALASAR³

- RESUMO: Previsões de resultados de jogos de futebol são de grande interesse por parte dos torcedores e imprensa, e têm sido o foco de várias pesquisas encontradas na literatura. Neste artigo, sob uma abordagem bayesiana, realizamos uma aplicação do modelo de regressão Poisson para prever os resultados do Campeonato Inglês de 2012–2013. Assumimos que os números de gols marcados pelos times em uma partida são independentes e seguem uma distribuição Poisson, em que a média reflete a força do ataque, defesa e casa. Antes do início de cada rodada do retorno, calculamos as probabilidades de vitória, empate e derrota dos times em cada uma das partidas e, por meio de um procedimento de simulação, obtivemos a probabilidade de um determinado time se classificar para a UEFA Champions League, sagrar-se campeão ou, também, de ser rebaixado. Todas as implementações computacionais foram realizadas utilizando os sistemas WinBUGS e R, por meio do pacote R2WinBUGS.
- PALAVRAS-CHAVE: Futebol; inferência bayesiana; previsão; regressão de Poisson; simulação.

¹Universidade de São Paulo - USP, Instituto de Ciências Matemáticas e de Computação, CEP: 13566-590, São Carlos, São Paulo, Brasil. E-mail: *cirodesc@gmail.com*; *suzuki@icmc.usp.br*; *louzada@icmc.usp.br*

²Universidade Federal de Mato Grosso do Sul - UFMS, Instituto de Matemática, CEP: 79070-900, Campo Grande, Mato Grosso do Sul, Brasil. E-mail: *erlandson.saraiva@ufms.br*

³Universidade Federal de São Carlos - UFSCar, Departamento de Estatística, Rod. Washington Luiz, km 235, CEP: 13565-905, São Carlos, São Paulo, Brasil. E-mail: *luis.salasar@gmail.com*

1 Introdução

O futebol é um esporte coletivo cuja paixão de muitos o torna o esporte mais popular do mundo. Em um determinado campeonato, nem sempre podemos credenciar o melhor time como sendo o vencedor de uma partida ou do torneio, por isso esta emoção é passada para os jogadores e, principalmente, para os torcedores.

Na literatura, poucos são os artigos que dizem respeito às previsões de placares da Copa do Mundo, tais como Dyte & Clarke (2000); Suzuki *et al.* (2009); Bastos & da Rosa (2013). O trabalho de Dyte & Clarke (2000) propôs um modelo de regressão de Poisson log-linear que considera o ranking da FIFA como covariável. Uma abordagem bayesiana para prever os resultados das partidas, utilizando a opinião dos especialistas e o ranking da FIFA como informações *a priori*, foi proposta por Suzuki *et al.* (2009). Já Bastos & da Rosa (2013) apresentam uma metodologia bayesiana do modelo Poisson-Gama para a previsão dos jogos da Copa do Mundo de 2010.

No entanto, vários são os artigos direcionados para a previsão de resultados de jogos de futebol aplicados às ligas nacionais. O modelo de Poisson foi utilizado por Keller (1994) para modelar os números de gols marcados por Inglaterra, Irlanda, Escócia e País de Gales no Campeonato Internacional Britânico (1883-1980) e por Louzada *et al.* (2014) para realizar previsões dos jogos do Campeonato Inglês de 2008-2009. Já Karlis & Ntzoufras (2009) modelam a diferença de gols entre as equipes mandante e visitante considerando a distribuição de Skellam.

Em geral, os times mandantes possuem uma taxa de vitória maior que a dos times visitantes (ver Pollard (1986) e Clarke & Norman (1995)). Na literatura, encontramos alguns trabalhos que levam em consideração na modelagem a vantagem de se jogar em casa, denominada efeito casa, ou fator casa. Por exemplo, o trabalho de Lee (1997) modela o número de gols por meio da distribuição de Poisson, em que a média de gols reflete o potencial ofensivo da equipe mandante, a capacidade defensiva da equipe visitante e o possível efeito do fator casa.

Tomando outra abordagem, as probabilidades de vitória, empate e derrota dos jogos do Campeonato Brasileiro da série A de 2006 são obtidas diretamente por Brillinger (2008), que considera um modelo de regressão trinomial.

Neste artigo, sob uma abordagem bayesiana, realizamos uma aplicação do modelo de regressão Poisson para prever os resultados do Campeonato Inglês de 2012-2013. O campeonato da primeira divisão do futebol inglês, também conhecido por *Premier League* ou simplesmente Campeonato Inglês, é a principal competição de futebol do país. O Campeonato Inglês de 2012-2013 foi disputado por 20 times no sistema de pontos corridos, atribuindo-se 3 pontos para cada vitória, 1 ponto para um empate e nenhum ponto para uma derrota. No final, o time que somar o maior número de pontos se consagra campeão, garantindo, automaticamente também, vaga para a Liga dos Campeões da UEFA, juntamente com o segundo e terceiro colocados. Já o quarto colocado disputa a fase eliminatória da Liga dos Campeões da UEFA, o quinto colocado tem vaga para a Liga Europa da UEFA e os três piores times são rebaixados para disputar a segunda divisão do futebol inglês.

Na modelagem, assumimos que os números de gols marcados pelos times em uma partida são independentes e seguem uma distribuição Poisson, em que a média reflete a força do ataque, da defesa e da casa. Antes do início de cada rodada do retorno, calculamos as probabilidades de vitória, empate e derrota dos times em cada uma das partidas e, por meio de um procedimento de simulação, obtemos a probabilidade de um determinado time se classificar para a Liga dos Campeões da UEFA, sagrar-se campeão, bem como de ser rebaixado para a segunda divisão.

Como nosso conjunto de dados, utilizamos os resultados do primeiro turno do Campeonato Inglês 2012–2013 (as 190 primeiras partidas das 19 primeiras rodadas disputadas) para realizarmos as previsões do retorno (os 190 jogos restantes referentes às rodadas 20 a 38). As primeiras 19 rodadas foram escolhidas como conjunto de treinamento pelo fato de ter observado uma partida de cada equipe contra os demais adversários.

O restante do artigo está organizado da seguinte maneira. Na Seção 2, apresentamos o modelo de regressão Poisson. Na Seção 3, mostramos as expressões utilizadas para calcular a previsão do resultado de uma determinada partida. Na Seção 4, apresentamos a aplicação do modelo na previsão de resultados dos jogos do campeonato Inglês de 2012-2013. Finalizamos o artigo com as considerações finais na Seção 5.

2 Modelo

Considere um campeonato de futebol com $n+1$ times. Cada time joga $2n$ vezes, sendo n vezes em casa e n vezes fora. O número total de jogos do campeonato é $N = n(n+1)$. Por cada resultado: vitória, empate e derrota, cada time recebe 3, 1 e 0 pontos, respectivamente. Ao final dos N jogos, o time que possui a maior quantidade de pontos é declarado campeão.

Considere X_{mi} e X_{vi} o número de gols marcados, respectivamente, pelo time mandante e visitante no i -ésimo jogo, para $i = 1, \dots, n$. Assuma que,

$$X_{mi} \sim \text{Poisson}(\lambda_{mi}) \quad \text{e} \quad X_{vi} \sim \text{Poisson}(\lambda_{vi})$$

com o número esperado de gols sendo dado pelo seguinte modelo de regressão:

$$\log(\lambda_{mi}) = \mu + \alpha_c + \alpha_{am} + \alpha_{dv} \quad \text{e} \quad \log(\lambda_{vi}) = \mu + \alpha_{av} + \alpha_{dm}, \quad (1)$$

em que μ é um parâmetro que representa o número médio de gols em uma partida, α_c é um parâmetro referente ao efeito casa, $(\alpha_{am}, \alpha_{dm})$ e $(\alpha_{av}, \alpha_{dv})$ são os parâmetros referentes à força de ataque e qualidade da defesa do time mandante e visitante, respectivamente, para $i = 1, \dots, n$. Note que, neste modelo, assumimos que um time com “boa” defesa apresenta α_{dt} negativo, $t = m, v$; pois isto diminui o valor do número esperado de gols do time adversário.

Assim, definindo n_r como sendo o número de jogos realizados pelos times m e v até a r -ésima rodada do campeonato e $\mathbf{x}_t = (x_{t1}, \dots, x_{tn_r})$ o vetor contendo

o número de gols marcado pelo time t , $t = m, v$, em que, x_{ti} é o número de gols marcado no i -ésimo jogo, então a função log-verossimilhança é dada por:

$$l(\boldsymbol{\alpha}_m, \boldsymbol{\alpha}_d | \mathbf{x}_m, \mathbf{x}_v) \propto \sum_{t \in \{m, v\}} \sum_{i=1}^{n_r} [x_{ti} \log(\lambda_{ti}) - \lambda_{ti}]$$

para $i = 1, \dots, n_r$, em que, $\log(\lambda_{ti})$ é dado em (1), $\boldsymbol{\alpha}_m = (\mu, \alpha_c, \alpha_{am}, \alpha_{dv})$ e $\boldsymbol{\alpha}_v = (\mu, \alpha_{av}, \alpha_{dm})$ são os parâmetros associados ao número esperado de gols dos times mandante e visitante, respectivamente.

Para evitar o problema de identificabilidade, é necessário impor algumas restrições sobre os parâmetros de ataque e defesa do modelo da seguinte forma (ver, por exemplos, Baio and Blangiardo (2010) e Karlis & Ntzoufras (2003)):

$$\sum_{m=1}^{n+1} \alpha_{am} = 0 \quad \text{e} \quad \sum_{m=1}^{n+1} \alpha_{dm} = 0, \quad (2)$$

i.e., a soma dos efeitos de ataque e defesa de todos os $n + 1$ times é igual a zero.

Para a análise bayesiana, assumimos que as distribuições *a priori* para $\boldsymbol{\alpha}_m$ e $\boldsymbol{\alpha}_v$ são mutuamente independentes, i.e., $\pi(\boldsymbol{\alpha}_m, \boldsymbol{\alpha}_v) = \pi(\boldsymbol{\alpha}_m)\pi(\boldsymbol{\alpha}_v)$, e que $\pi(\boldsymbol{\alpha}_m) = \pi(\mu)\pi(\alpha_c)\pi(\alpha_{am})\pi(\alpha_{dv})$ e $\pi(\boldsymbol{\alpha}_v) = \pi(\mu)\pi(\alpha_{av})\pi(\alpha_{dm})$. Consideramos as seguintes distribuições *a priori*, $\mu \sim \mathcal{N}(0, g^2)$, $\alpha_c \sim \mathcal{N}(0, g^2)$, $\alpha_{at} \sim \mathcal{N}(0, g^2)$ e $\alpha_{dt} \sim \mathcal{N}(0, g^2)$, para $t = m, v$, em que $\mathcal{N}(0, g^2)$ denota a distribuição Normal com média zero e variância g^2 . Fixamos o hiperparâmetro g^2 igual a 100 para obtermos uma distribuição *a priori* não informativa. Um estudo de sensibilidade foi realizado escolhendo-se outros valores para este hiperparâmetro e foram observadas estimativas *a posteriori* muito próximas.

Como a distribuição *a posteriori* conjunta para os parâmetros de interesse não possuem forma fechada conhecida, estimamos os parâmetros via métodos de Monte Carlo em Cadeias de Markov (MCMC). Todas as implementações computacionais foram realizadas utilizando os sistemas WinBUGS (Spiegelhalter *et al.*, 2003) e R (R CORE TEAM, 2016), considerando o pacote R2WinBUGS (Gelman *et al.*, 2006).

Sejam $\tilde{\boldsymbol{\alpha}}_m = (\tilde{\mu}, \tilde{\alpha}_c, \tilde{\alpha}_{am}, \tilde{\alpha}_{dv})$ e $\tilde{\boldsymbol{\alpha}}_v = (\tilde{\mu}, \tilde{\alpha}_{av}, \tilde{\alpha}_{dm})$ as estimativas de $\boldsymbol{\alpha}_m$ e $\boldsymbol{\alpha}_v$, respectivamente. Estas estimativas são dadas pelas médias dos valores MCMC gerados. Condicional $\tilde{\boldsymbol{\alpha}}_m$ e $\tilde{\boldsymbol{\alpha}}_v$, calculamos as probabilidades de vitória, empate e derrota para cada time na próxima rodada do campeonato.

3 Métodos

Nesta seção, mostramos as expressões utilizadas para calcular a previsão de uma determinada partida. Além disso, descrevemos o procedimento de simulação para calcular as diversas probabilidades de interesses, tais como: de campeão, de rebaixamento e de classificações para as competições europeias. Analogamente a Suzuki *et al.* (2009) e Bastos & da Rosa (2013), utilizamos duas métricas para comparação das qualidades das previsões: a medida de Finetti (de Finetti, 1972) e o número de acertos.

3.1 Previsão para o resultado de uma partida

Em uma partida de futebol, é declarado vencedor a equipe que marcar mais gols. O empate ocorre se as equipes não marcam gols ou marcam o mesmo número de gols.

A probabilidade de vitória (P_V), empate (P_E) e derrota (P_D) para o time mandante m contra o time visitante v , no i -ésimo jogo do campeonato, é obtida utilizando as expressões:

$$P_V = P(X_{mi} > X_{vi} | \tilde{\alpha}_m, \tilde{\alpha}_v) = \sum_{g=1}^{\infty} \sum_{u=0}^{g-1} P(X_{mi} = g | \tilde{\alpha}_m) P(X_{vi} = u | \tilde{\alpha}_v); \quad (3)$$

$$P_E = P(X_{mi} = X_{vi} | \tilde{\alpha}_m, \tilde{\alpha}_v) = \sum_{g=0}^{\infty} P(X_{mi} = g | \tilde{\alpha}_m) P(X_{vi} = g | \tilde{\alpha}_v); \quad (4)$$

$$P_D = P(X_{mi} < X_{vi} | \tilde{\alpha}_m, \tilde{\alpha}_v) = \sum_{u=1}^{\infty} \sum_{g=0}^{u-1} P(X_{mi} = g | \tilde{\alpha}_m) P(X_{vi} = u | \tilde{\alpha}_v). \quad (5)$$

3.2 Previsão para Classificação Final do Campeonato

Antes de cada uma das rodadas do segundo turno (rodadas 20 a 38), a previsão para a classificação final foi predita utilizando o seguinte algoritmo.

Algoritmo. Seja T_t o número de pontos de um time t ao final da primeira fase do campeonato, para $t = 1, \dots, n + 1$.

- (i) Para o i -ésimo jogo da segunda fase do campeonato, $i = (n(n + 1)/2) + 1, \dots, N$, faça:
 - (a) Obtenhas as estimativas $(\tilde{\alpha}_m, \tilde{\alpha}_v)$;
 - (b) Gere $X_{mi} \sim Poisson(\lambda_{mi})$ e $X_{vi} \sim Poisson(\lambda_{vi})$, onde λ_{mi} e λ_{vi} são dados em (1);
 - (c) Se $X_{mi} > X_{vi}$, faça $T_m = T_m + 3$ e $T_v = T_v$; se $X_{mi} = X_{vi}$, faça $T_m = T_m + 1$ e $T_v = T_v + 1$; se $X_{mi} < X_{vi}$, faça $T_m = T_m$ e $T_v = T_v + 3$;
 - (d) Terminado todos os jogos da segunda fase, um time A é declarado campeão se $T_a = \max_{1 \leq t \leq n} T_t$.
- (ii) Repita o passo (i), l vezes, for $l = 1, \dots, L$. Isto é, simulamos L campeonatos.

Fixamos $L = 1.000$. A partir daí, podemos obter a estimativa da probabilidade de um determinado time ser campeão e de ser rebaixado da seguinte forma:

$$P_{\text{campeão}} = \frac{\#(\text{time terminou em primeiro lugar})}{L} \quad (6)$$

e

$$P_{\text{rebaixamento}} = \frac{\#(\text{time terminou entre os 4 \u00faltimos colocados})}{L}, \quad (7)$$

em que, $\#$ refere-se ao n\u00famero de vezes obtidos nos L campeonatos simulados.

3.3 Qualidade das Previs\u00f5es

Uma m\u00e9trica para a verifica\u00e7\u00e3o da qualidade das previs\u00f5es dos resultados de uma partida de futebol \u00e9 a medida de Finetti (de Finetti, 1972), ver por exemplo Suzuki *et al.* (2009); Bastos & da Rosa (2013) e Louzada *et al.* (2014). Esta medida consiste na considera\u00e7\u00e3o de um simplex contido em \mathbf{R}^3 como representa\u00e7\u00e3o geom\u00e9trica do conjunto das poss\u00edveis previs\u00f5es probabil\u00edsticas. Assim, os v\u00e9rtices desse simplex correspondem \u00e0s ocorr\u00eancias dos resultados e os demais pontos a todas as outras poss\u00edveis previs\u00f5es. Formalmente,

$$\mathbb{S} = \{(P_V, P_E, P_D) \in \mathbf{R}^3 : P_V + P_E + P_D = 1, P_V \geq 0, P_E \geq 0, P_D \geq 0\},$$

em que, P_V denota a probabilidade de vit\u00f3ria do time mandante, P_E a probabilidade de empate e P_D a probabilidade de derrota do time mandante. A medida de dist\u00e2ncia de Finetti corresponde ao quadrado da dist\u00e2ncia euclidiana entre o ponto correspondente da probabilidade prevista e o v\u00e9rtice correspondente ao resultado efetivamente observado. Para mais de uma previs\u00e3o, pode-se construir um \u00edndice dado pela m\u00e9dia aritm\u00e9tica das dist\u00e2ncias de Finetti chamado Medida de Finetti.

Na aplica\u00e7\u00e3o no futebol, associam-se os v\u00e9rtices $(1, 0, 0)$, $(0, 1, 0)$ e $(0, 0, 1)$ \u00e0 vit\u00f3ria da equipe mandante, ao empate e \u00e0 derrota da equipe mandante, respectivamente. Ao vetor de probabilidades atribu\u00eddas para uma determinada partida, associa-se o ponto $(P_V, P_E, P_D) \in \mathbb{S}$. Assim, a dist\u00e2ncia de Finetti \u00e9:

$$(P_V - 1)^2 + (P_E - 0)^2 + (P_D - 0)^2 \text{ se a equipe mandante vencer a partida;}$$

$$(P_V - 0)^2 + (P_E - 1)^2 + (P_D - 0)^2 \text{ se a partida terminar empatada;}$$

$$(P_V - 0)^2 + (P_E - 0)^2 + (P_D - 1)^2 \text{ se a equipe mandante perder a partida.}$$

Por exemplo, se a previs\u00e3o for $(0, 2, 0, 65, 0, 15)$ e o resultado observado \u00e9 um empate $(0, 1, 0)$, ent\u00e3o a dist\u00e2ncia de Finetti \u00e9 $(0, 2 - 0)^2 + (0, 65 - 1)^2 + (0, 15 - 0)^2 = 0, 185$. Em termos de previs\u00e3o, temos que um padr\u00e3o comumente utilizado \u00e9 o da atribui\u00e7\u00e3o equiprov\u00e1vel de probabilidades $(P_V = P_E = P_D = 1/3)$, que, em outras palavras, atribui chances iguais a cada resultado em cada jogo. Para essa atribui\u00e7\u00e3o, a medida de Finetti \u00e9 igual a $(1/3 - 1)^2 + (1/3 - 0)^2 + (1/3 - 0)^2 = 2/3$. Assim, podemos considerar m\u00e9todos de previs\u00f5es de qualidade minimamente aceit\u00e1veis aqueles que apresentam medidas de Finetti menores que $2/3$ e de m\u00e1-qualidade maiores que $2/3$.

Com rela\u00e7\u00e3o \u00e0 quantidade de n\u00famero de acertos, consideramos que a previs\u00e3o acerta o resultado de um determinado jogo quando o evento com maior probabilidade estimada \u00e9 observado. Por exemplo, se em um determinado jogo o time mandante vencer e a probabilidade estimada para a vit\u00f3ria do time mandante

for maior que as probabilidades de empate e derrota, consideramos um acerto para esta previsão, caso contrário, um erro. O mesmo vale para a ocorrência de empate e vitória do time visitante. No entanto, vale ressaltar que não é correto associar a eventos com alta probabilidade uma certeza de ocorrência ou a eventos com baixa probabilidade uma certeza de não ocorrência. Aqui, o objetivo não é afirmar que uma dada previsão é correta ou não mas, sim, construir uma métrica para o conjunto das previsões.

4 Resultados

Nesta seção, apresentamos os resultados obtidos da aplicação no Campeonato Inglês de 2012–2013. Inicialmente, realizamos uma breve análise exploratória dos resultados dos jogos. Em seguida, são apresentados tanto os resultados para as previsões das partidas de uma determinada rodada, quanto para o campeonato inteiro. Para as previsões das partidas de uma rodada, utilizamos os resultados dos jogos obtidos até a rodada anterior. Assim, para prever a rodada 31, por exemplo, foram usados os resultados dos jogos pertencentes às rodadas anteriores à rodada 31. Também, antes de cada rodada do retorno (2º Turno), foram simulados 1.000 campeonatos para obter as probabilidades de cada time ocupar uma determinada posição.

4.1 Análise descritiva

Nesta seção, realizamos uma análise descritiva dos dados do Campeonato Inglês da temporada 2012/2013. A Tabela 1 indica o número de jogos, postos entre parênteses, que cada equipe fez como mandante e visitante e o número de gols marcados, considerando os jogos das 19 primeiras rodadas (1º Turno), os 19 jogos do retorno (2º Turno) além de considerar também todas as 38 rodadas (Geral). Como podemos observar, o Arsenal teve o melhor ataque como mandante (47 gols) enquanto que o Manchester United foi o melhor do campeonato (86 gols), além de ser o time que mais fez gols como visitante (41 gols).

A Tabela 2 refere-se ao total de todos os placares que ocorreram ao longo do campeonato, em que GMM significa os gols marcados pelas equipes mandantes e GMV os gols marcados pelas equipes visitantes. Pode-se notar que os resultados que aconteceram com maior frequência foram empate por 1x1 e vitórias dos mandantes pelos placares de 1x0 e 2x1.

A soma dos valores de cada coluna da Tabela 2 dá o número de jogos que o time mandante marcou x_m gols, $x_m \in \{1, 2, 3, 4+\}$. Por exemplo, a soma dos valores da segunda coluna dá 120; ou seja, em 120 jogos o time mandante marcou 2 gols. De forma análoga, a soma dos valores de cada linha dá o número de jogos que o time visitante marcou x_v gols, $x_v \in \{1, 2, 3, 4+\}$. Por exemplo, somando os valores da terceira linha obtemos o valor 37; ou seja, em 37 jogos o time visitante marcou 3 gols. A diagonal principal desta Tabela dá o número de jogos que terminou em empate. O total de empates é 109 (28,68%). A soma dos valores das diagonais superior e inferior dão as quantidades de jogos vencidos pelos times mandantes e

visitantes, respectivamente. O total de vitórias dos times mandantes foram 166 (43.68%) e dos times visitantes foram 105 (27.64%).

Tabela 1 - Número de gols marcados pelos times no campeonato

Times	1º Turno		2º Turno		Geral	
	(Mandante) Gols	(Visitante) Gols	(Mandante) Gols	(Visitante) Gols	(Mandante) Gols	(Visitante) Gols
Arsenal	(9) 23	(10) 14	(10) 24	(9) 11	(19) 47	(19) 25
Aston Villa	(9) 8	(10) 7	(10) 15	(9) 17	(19) 23	(19) 24
Chelsea	(10) 24	(9) 15	(9) 17	(10) 19	(19) 41	(19) 34
Everton	(9) 16	(10) 16	(10) 17	(9) 8	(19) 33	(19) 22
Fulham	(9) 16	(10) 13	(10) 12	(9) 9	(19) 28	(19) 22
Liverpool	(10) 14	(9) 14	(9) 19	(10) 24	(19) 33	(19) 38
Manchester City	(10) 22	(9) 12	(9) 19	(10) 13	(19) 41	(19) 25
Manchester United	(9) 26	(10) 22	(10) 19	(9) 19	(19) 45	(19) 41
NewCastle United	(10) 12	(9) 11	(9) 12	(10) 10	(19) 24	(19) 21
Norwich City	(9) 10	(10) 10	(10) 15	(9) 6	(19) 25	(19) 16
Queens Park Rangers	(9) 8	(10) 8	(10) 5	(9) 9	(19) 13	(19) 17
Reading	(9) 14	(10) 7	(10) 9	(9) 10	(19) 23	(19) 20
Southampton	(10) 14	(9) 11	(9) 12	(10) 12	(19) 26	(19) 23
Stoke City	(9) 11	(10) 7	(10) 10	(9) 6	(19) 21	(19) 13
Sunderland	(9) 10	(10) 10	(10) 10	(9) 11	(19) 20	(19) 21
Swansea City	(10) 17	(9) 10	(9) 11	(10) 9	(19) 28	(19) 19
Tottenham Hotspur	(10) 14	(9) 20	(9) 15	(10) 17	(19) 29	(19) 37
West Bromwich Albion	(10) 16	(9) 12	(9) 16	(10) 9	(19) 32	(19) 21
West Ham United	(10) 17	(9) 6	(9) 17	(10) 5	(19) 34	(19) 11
Wigan	(10) 14	(9) 6	(9) 12	(10) 15	(19) 26	(19) 21

Tabela 2 - Total de placares do campeonato

GMV	GMM					Total
	0	1	2	3	4+	
0	35	41	18	12	10	116
1	20	42	41	18	10	131
2	13	27	27	09	05	81
3	10	10	11	04	02	37
4+	06	00	05	03	01	15
Total	84	120	102	46	28	380

4.2 Previsão para o resultado de uma partida

Calculamos as probabilidades de vitória, empate e derrota (ver Subseção 3.1) das partidas de todas as rodadas do segundo turno, ou seja, antes de cada uma das 19 rodadas restantes (rodadas 20 a 38). A medida de Finetti (de Finetti, 1972) associada a essas previsões é igual a 0,595.

Para obtenção das estimativas consideramos 50.000 iterações das quais as 10.000 primeiras iterações foram descartadas como *burn in*. Após o *burn in* selecionamos um valor a cada 20 (salto = 20) para inferência.

Dessa forma, obtemos uma amostra de tamanho 2.000, sobre a qual a inferência *a posteriori* foi feita.

Na Tabela 3, apresentamos a previsão para cada uma das partidas da rodada 31 juntamente com o placar observado, a medida de Finetti e se ocorreu acerto. Podemos observar que as probabilidades de vitória são maiores quando o time é mandante e, caso o time possua um ataque forte, como é o caso do Arsenal, a sua

probabilidade de ganhar a partida aumenta e, neste caso, é de 0,566. Isso pode ser considerado uma grande diferença, pois enfrentando um time mais fraco como o Reading, este tem apenas 0,217 de probabilidade de vitória. O resultado observado neste jogo foi a vitória do time mandante.

A proporção de predição correta para esta rodada é de 80%.

Tabela 3 - Probabilidades de resultados, medida de Finetti para rodada 31

Mandante	Visitante	Probabilidades			Placar	de Finetti	Acerto
		Vitória	Empate	Derrota			
Arsenal	Reading	0,566	0,217	0,217	4 - 1	0,282	Sim
Aston	Liverpool	0,303	0,257	0,439	1 - 2	0,473	Sim
Everton	Stoke	0,636	0,203	0,162	1 - 0	0,200	Sim
Fulham	Queens Park Rangers	0,433	0,260	0,307	3 - 2	0,483	Sim
Manchester City	Newcastle	0,570	0,239	0,191	4 - 0	0,278	Sim
Southampton	Chelsea	0,383	0,262	0,355	2 - 1	0,575	Sim
Sunderland	Manchester United	0,457	0,240	0,303	0 - 1	0,752	Não
Swansea	Tottenham	0,397	0,219	0,384	1 - 2	0,585	Não
WestHam	WestBrom	0,538	0,224	0,238	3 - 1	0,320	Sim
Wigan	Norwich	0,511	0,249	0,241	1 - 0	0,359	Sim

No Apêndice A, apresentamos, como ilustração, as estimativas dos parâmetros de interesse e os gráficos das autocorrelações estimadas para o melhor time do campeonato (Manchester United) e seu adversário nas rodadas 20 e 30.

4.3 Previsão para Classificação Final do Campeonato

Para obter as outras previsões de interesses, tais como a probabilidade de ser campeão e a probabilidade de ser rebaixado, realizamos uma simulação do campeonato inteiro. Na Tabela 4, apresentamos as probabilidades de cada uma das sete melhores equipes de se tornarem campeãs da competição antes de cada rodada do segundo turno. Observamos, com exceção da rodada 21, um amplo favoritismo do Manchester United.

Na Tabela 5, apresentamos as probabilidades de conquista de uma vaga para a competição europeia de cada uma das sete melhores equipes. Observamos que os times Manchester United e Manchester City conquistaram suas vagas antecipadas para a disputa da Liga dos Campeões da UEFA. As outras duas vagas foram disputadas entre três equipes de Londres: Arsenal, Chelsea e Tottenham, e apenas na última rodada as vagas foram preenchidas pelos times que tinham maiores probabilidades, que foram o Arsenal e o Chelsea. O Tottenham acabou conquistando a vaga para a Liga Europa da UEFA, juntamente com a equipe do Swansea (vaga conquistada como campeão da Copa da Liga Inglesa 2012-2013).

Na Tabela 6, apresentamos as probabilidades de rebaixamento das oito piores equipes. Neste campeonato, ao término da trigésima oitava rodada, os últimos três colocados caem para a segunda divisão. As equipes rebaixadas foram o Wigan, Reading e Queens Park Rangers, que, a partir da rodada 33, eram os times com maiores probabilidades descenso.

A Figura 1 apresenta o diagrama em caixa da pontuação final das equipes antes da rodada 32. Por meio das simulações realizadas, podemos observar a dispersão

Tabela 4 - Probabilidade de conquista do título para cada uma das sete melhores equipes do campeonato

Rodada	Manchester United	Manchester City	Chelsea	Arsenal	Tottenham	Everton	Liverpool
20	0,429	0,208	0,012	0,006	0,087	0,172	0,000
21	0,134	0,516	0,161	0,109	0,005	0,001	0,060
22	0,848	0,053	0,002	0,028	0,051	0,002	0,003
23	0,561	0,382	0,024	0,001	0,023	0,003	0,001
24	0,851	0,141	0,002	0,001	0,001	0,003	0,000
25	0,581	0,414	0,004	0,000	0,000	0,001	0,000
26	0,958	0,038	0,002	0,000	0,000	0,002	0,000
27	0,906	0,089	0,002	0,000	0,002	0,000	0,000
28	0,863	0,111	0,014	0,000	0,012	0,000	0,000
29	0,903	0,094	0,000	0,000	0,003	0,000	0,000
30	0,969	0,030	0,001	0,000	0,000	0,000	0,000
31	0,964	0,033	0,003	0,000	0,000	0,000	0,000
32	0,998	0,001	0,001	0,000	0,000	0,000	0,000
33	0,979	0,021	0,000	0,000	0,000	0,000	0,000
34	0,997	0,003	0,000	0,000	0,000	0,000	0,000
35	1,000	0,000	0,000	0,000	0,000	0,000	0,000
36	1,000	0,000	0,000	0,000	0,000	0,000	0,000
37	1,000	0,000	0,000	0,000	0,000	0,000	0,000
38	1,000	0,000	0,000	0,000	0,000	0,000	0,000

Tabela 5 - Probabilidade de conquista de uma vaga para competições europeias para cada uma das sete melhores equipes do campeonato

Rodada	Manchester United	Manchester City	Chelsea	Arsenal	Tottenham	Everton	Liverpool
20	0,883	0,736	0,162	0,074	0,520	0,689	0,019
21	0,733	0,920	0,704	0,638	0,132	0,010	0,526
22	0,997	0,681	0,125	0,527	0,701	0,143	0,095
23	0,987	0,969	0,589	0,135	0,589	0,118	0,194
24	0,999	0,970	0,348	0,278	0,410	0,620	0,091
25	1,000	0,998	0,622	0,435	0,304	0,338	0,226
26	1,000	0,945	0,728	0,214	0,306	0,644	0,103
27	1,000	0,964	0,571	0,475	0,466	0,078	0,075
28	1,000	0,983	0,879	0,207	0,842	0,031	0,051
29	1,000	0,998	0,809	0,101	0,945	0,098	0,039
30	1,000	0,975	0,935	0,186	0,842	0,037	0,018
31	1,000	0,996	0,979	0,259	0,632	0,091	0,036
32	1,000	0,993	0,805	0,555	0,504	0,071	0,071
33	1,000	1,000	0,804	0,473	0,630	0,093	0,000
34	1,000	1,000	0,772	0,508	0,587	0,130	0,003
35	1,000	1,000	0,796	0,927	0,269	0,008	0,000
36	1,000	1,000	0,926	0,642	0,380	0,052	0,000
37	1,000	1,000	0,896	0,770	0,334	0,000	0,000
38	1,000	1,000	1,000	1,000	0,000	0,000	0,000

dos pontos, as maiores e as menores pontuações. Claramente, a pontuação do Manchester United sobressai em relação às demais equipes.

A Tabela 7 mostra as probabilidades de cada uma das equipes participantes do campeonato terminarem em cada uma das vinte possíveis posições, segundo as simulações de seus desempenhos até o final do campeonato, considerando as estimativas dos parâmetros utilizando os dados até a trigésima rodada. Também são apresentadas a média e o desvio padrão (DP) das posições de cada um dos

Tabela 6 - Probabilidade de rebaixamento para cada uma das oito piores equipes do campeonato

Rodada	Stoke	Southampton	Aston	Newcastle	Sunderland	Wigan	Reading	Queens Park Rangers
20	0,008	0,611	0,115	0,211	0,044	0,083	0,788	0,765
21	0,130	0,147	0,459	0,031	0,011	0,080	0,322	0,988
22	0,000	0,179	0,334	0,344	0,138	0,795	0,206	0,982
23	0,010	0,171	0,547	0,106	0,052	0,530	0,477	0,708
24	0,063	0,083	0,742	0,197	0,001	0,299	0,378	0,831
25	0,011	0,020	0,504	0,144	0,062	0,471	0,809	0,669
26	0,135	0,101	0,702	0,089	0,000	0,813	0,311	0,411
27	0,000	0,164	0,547	0,054	0,078	0,814	0,249	0,956
28	0,000	0,061	0,306	0,144	0,327	0,241	0,805	0,904
29	0,012	0,255	0,474	0,029	0,029	0,391	0,831	0,941
30	0,073	0,062	0,321	0,009	0,113	0,839	0,929	0,622
31	0,011	0,304	0,134	0,105	0,117	0,445	0,864	0,868
32	0,045	0,124	0,336	0,072	0,199	0,225	0,911	0,944
33	0,056	0,036	0,221	0,052	0,245	0,414	0,996	0,894
34	0,059	0,009	0,104	0,029	0,058	0,714	0,987	0,989
35	0,006	0,001	0,183	0,015	0,005	0,782	1,000	0,999
36	0,001	0,000	0,107	0,015	0,023	0,825	1,000	1,000
37	0,000	0,004	0,003	0,138	0,035	0,777	1,000	1,000
38	0,000	0,000	0,000	0,000	0,000	1,000	1,000	1,000

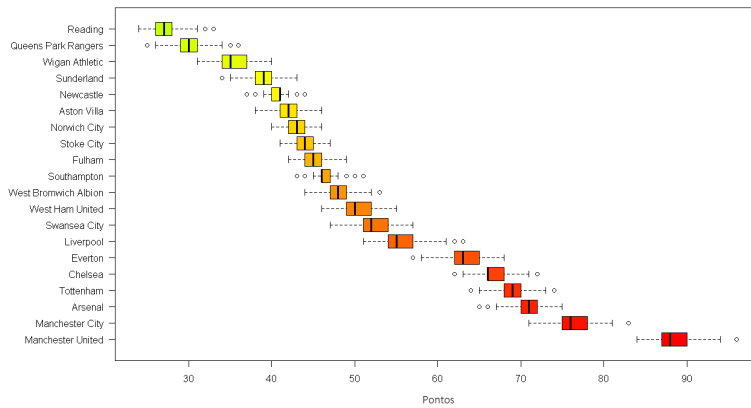


Figura 1 - Diagrama em caixa da pontuação final das equipes antes da rodada 32.

vinte times para os 1000 campeonatos simulados. Observamos que, nesta rodada, a probabilidade do Manchester United se tornar campeão é de 0,947, e do Manchester City de ser vice-campeão é de 0,845.

Tabela 7 - Probabilidade de cada equipe terminar em cada uma das 20 posições

Equipe	Média	DP	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	12°	13°	14°	15°	16°	17°	18°	19°	20°	
Manchester United	1,053	0,224	0,947	0,053	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Manchester City	2,068	0,460	0,953	0,845	0,085	0,015	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Chelsea	3,498	0,927	0,000	0,072	0,550	0,223	0,122	0,029	0,004	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Arsenal	4,700	1,024	0,000	0,005	0,111	0,302	0,390	0,152	0,033	0,007	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Tottenham	4,278	1,104	0,000	0,025	0,226	0,366	0,246	0,106	0,028	0,063	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Everton	6,098	1,112	0,000	0,000	0,017	0,073	0,161	0,374	0,305	0,059	0,011	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Liverpool	6,628	1,069	0,000	0,000	0,011	0,021	0,077	0,303	0,439	0,117	0,024	0,005	0,003	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
West Brom	8,852	1,194	0,000	0,000	0,000	0,000	0,002	0,031	0,143	0,485	0,201	0,081	0,036	0,016	0,003	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Swansea	10,244	1,737	0,000	0,000	0,000	0,000	0,000	0,001	0,019	0,116	0,239	0,245	0,172	0,103	0,053	0,028	0,018	0,005	0,001	0,000	0,000	0,000	0,000
West Ham	11,299	1,797	0,000	0,000	0,000	0,000	0,000	0,000	0,003	0,032	0,121	0,202	0,221	0,180	0,130	0,055	0,038	0,014	0,004	0,000	0,000	0,000	0,000
Norwich	13,323	2,048	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,023	0,054	0,125	0,151	0,180	0,167	0,139	0,100	0,043	0,013	0,002	0,000	0,000
Fulham	9,716	1,483	0,000	0,000	0,000	0,000	0,000	0,004	0,024	0,161	0,319	0,235	0,141	0,070	0,028	0,012	0,004	0,002	0,000	0,000	0,000	0,000	0,000
Stoke	13,436	2,073	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,005	0,021	0,061	0,088	0,168	0,154	0,188	0,152	0,094	0,054	0,008	0,007	0,000	0,000
Southampton	12,925	2,089	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,012	0,035	0,083	0,126	0,158	0,196	0,157	0,117	0,072	0,034	0,008	0,000	0,001	0,000
Aston Villa	17,017	1,688	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,004	0,010	0,019	0,048	0,088	0,158	0,240	0,247	0,150	0,035	0,000
New Castle	14,826	1,996	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,005	0,014	0,044	0,069	0,107	0,162	0,198	0,205	0,119	0,060	0,011	0,005	0,000
Sunderland	14,881	2,024	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,019	0,039	0,072	0,121	0,152	0,184	0,189	0,132	0,070	0,020	0,001	0,000
Wigan	17,642	1,475	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,003	0,008	0,023	0,045	0,104	0,242	0,256	0,252	0,066	0,000
Reading	19,375	0,955	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,002	0,003	0,011	0,035	0,108	0,233	0,608	0,000
Queens Park Rangers	18,641	1,246	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,005	0,013	0,046	0,096	0,230	0,325	0,284	0,000

Apresentamos na Figura 2, o gráfico de pontos do efeito de ataque *versus* efeito de defesa, de todas as equipes. Para a sua construção, consideramos todo o conjunto de dados, ou seja, todos os resultados após as 38 rodadas. Nota-se que as melhores equipes se destacam pelo ataque forte e pela defesa forte, enquanto que as equipes com defesa fraca sofrem no torneio.

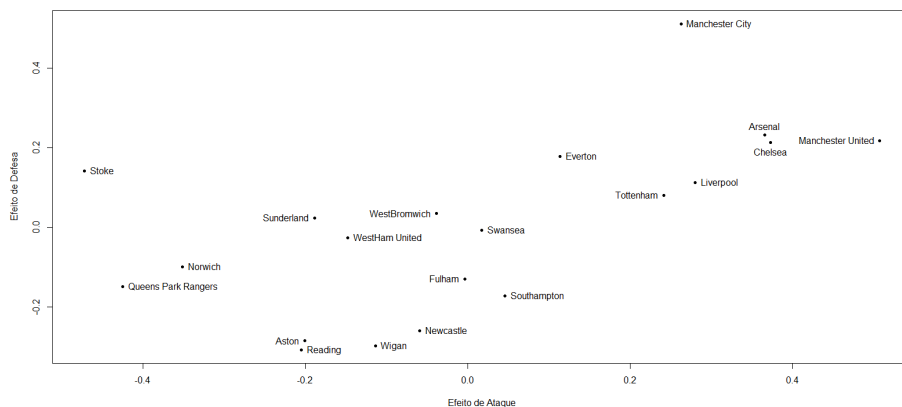


Figura 2 - Gráfico de pontos dos efeitos de ataque e defesa.

Na Figura 3 mostramos, para as quatro melhores e para as quatro piores equipes do campeonato, o comportamento do efeito de ataque (gráfico superior esquerdo e gráfico inferior esquerdo, respectivamente) e do efeito de defesa (gráfico superior direito e gráfico inferior direito, respectivamente), ao longo das rodadas 20 a 38.

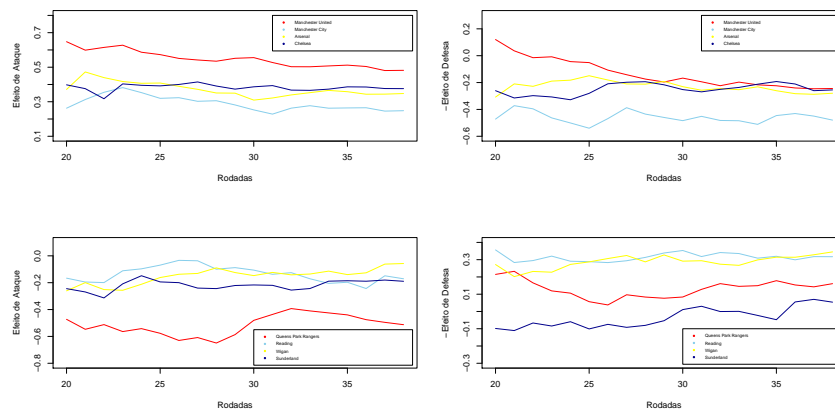


Figura 3 - Gráfico de linhas para o efeito de ataque e para o efeito de defesa.

A Tabela 8 apresenta a previsão da classificação final antes da trigésima rodada, em comparação com as reais posições. Para obter esta tabela, foram previstas 90 partidas, e observamos, já nesta rodada, uma ligeira vantagem do campeão, Manchester United, e os três times rebaixados nas últimas posições.

Tabela 8 - Previsão da classificação final antes da rodada 30 (Classificação final do campeonato)

Equipes	Pontos	Jogos	Vitórias	Empates	Derrotas	Gols Pro	Gols Contra	Saldo
Manchester United	92(89)	38	29(28)	5(5)	4(5)	88(86)	39(43)	49(43)
Manchester City	74(78)	38	22(23)	8(9)	8(6)	64(66)	39(34)	25(32)
Chelsea	71(75)	38	20(22)	11(9)	7(7)	75(75)	41(39)	34(36)
Arsenal	71(73)	38	20(21)	11(10)	7(7)	68(66)	36(37)	32(35)
Tottenham	65(72)	38	18(21)	11(9)	9(8)	65(66)	51(46)	14(20)
Everton	62(63)	38	16(16)	14(15)	8(7)	57(55)	42(40)	15(15)
Liverpool	56(61)	38	14(16)	14(13)	10(9)	69(71)	47(43)	22(28)
West Bromwich	48(49)	38	14(14)	6(7)	18(17)	48(53)	53(57)	-5(-4)
Swansea	49(46)	38	12(11)	13(13)	13(14)	50(47)	51(51)	-1(-4)
West Ham	49(46)	38	13(12)	10(10)	15(16)	45(45)	51(53)	-6(-8)
Norwich	44(44)	38	10(10)	14(14)	14(14)	38(41)	58(58)	-20(-17)
Fulham	42(43)	38	10(11)	12(10)	16(17)	49(50)	61(60)	-12(-10)
Stoke	42(42)	38	9(9)	15(15)	14(14)	35(34)	46(45)	-11(-11)
Southampton	44(41)	38	10(9)	14(14)	14(15)	50(49)	59(60)	-9(-11)
Aston	44(41)	38	11(10)	11(11)	16(17)	46(47)	66(69)	-20(-22)
NewCastle	37(41)	38	10(11)	7(8)	21(19)	44(45)	70(68)	-26(-23)
Sunderland	43(39)	38	11(9)	10(12)	17(17)	46(41)	55(54)	-9(-13)
Wigan	35(36)	38	9(9)	8(9)	21(20)	43(47)	67(73)	-24(-26)
Reading	32(28)	38	7(6)	11(10)	20(22)	48(43)	70(73)	-22(30)
Queens Park Rangers	30(25)	38	5(4)	15(13)	18(21)	31(30)	57(60)	-26(-30)

5 Considerações Finais

Neste artigo, sob uma abordagem bayesiana, realizamos uma aplicação do modelo de regressão de Poisson, em que a média reflete a força do ataque, da defesa e do mando de jogo. A metodologia apresentada tem uma boa capacidade preditiva, de acordo com a medida de Finetti e com a quantidade de acertos. Aliada a um procedimento de simulação, pode-se obter a estimativa da probabilidade de uma equipe ser campeã ou ser rebaixada, bem como a previsão do número de pontos a partir de uma rodada específica. Outros questionamentos de interesses também podem ser respondidos, tais como: Qual a probabilidade de classificação para a fase de grupos da Liga dos Campeões da UEFA 2013-2014? Qual equipe vai ter o melhor ataque? Qual equipe vai perder mais jogos? Qual equipe terá mais vitórias?, dentre outros. Além disso, a metodologia considerada pode ser facilmente utilizada para analisar as próximas temporadas do campeonato e ser adaptada para outros torneios com diferentes formas de disputa.

Todas as implementações computacionais foram realizadas utilizando os sistemas WinBUGS (Spiegelhalter *et al.*, 2003) e R (R CORE TEAM, 2016) considerando do pacote R2WinBUGS (Gelman *et al.*, 2006). Os códigos WinBUGS estão apresentados no Apêndice C. Os resultados dos jogos, como também outras informações adicionais, podem ser obtidas no site oficial da competição (<http://www.premierleague.com>).

Como trabalho futuro, pretendemos incorporar outras covariáveis no modelo, tais como: condição atmosférica, esquema tático, contusões, crise e suspensões. Além disso, assumindo que o resultado de uma partida é um vetor aleatório bivariado, cujas componentes representam a contagem do número de gols marcados por cada uma das equipes, uma possibilidade seria a de modelar o problema a partir da estrutura de regressão de um modelo bivariado como, por exemplo, a distribuição Poisson bivariada de Holgate (Holgate , 1964).

Agradecimentos

Agradecemos ao editor e aos revisores pela cuidadosa leitura, comentários e sugestões que contribuíram para uma melhora do artigo. O autor Erlandson F. Saraiva agradece o apoio do CNPq. O autor Francisco Louzada agradece o apoio do CNPq e da FAPESP.

FILHO, C. A. O.; SUZUKI, A. K., LOUZADA, F., SARAIVA, E. F., SALASAR, L. E. B. A Bayesian approach to predicting football match scores: An application to the English Premier League. *Rev.Bras. Biom.*, Lavras, v.35, n.1, p.76-97, 2017.

■ **ABSTRACT:** *The prediction of outcomes of football matches is of great interest to fans and the press, and has been the focus of several studies in the literature. In this manuscript, we carry out an application of the Poisson regression model for the prediction of outcomes of soccer games of the 2012–2013 English Premier League under a Bayesian approach. We assume that the number of goals scored by each team in a match are independent and follow a Poisson distribution, whose average reflects the strength of the attack, defense and home advantage parameter. Before the start of each round of the second round, we calculated the win, draw and loss probabilities for each match and, through a simulation procedure, we have obtained the probability of a team qualifying for the UEFA Champions League, being crowned champion or relegated to the second division is obtained. All computer implementations were performed using WinBUGS and R systems through the R2WinBUGS package.*

■ **KEYWORDS:** *Football; bayesian inference; prediction; Poisson regression; simulation.*

Referências

BAIO, G.; BLANGIARDO, M. Bayesian hierarchical model for the prediction of football result. *UCL Discovery*, v.37, p.253-264, 2010.

BASTOS, L. S.; da ROSA, J. M. C. Predicting probabilities for the 2010 FIFA world cup games using a Poisson-Gamma model. *Journal of Applied Statistics*, v.40, n.7, p.1533-1544, 2013.

- BRILLINGER, D. R. Modelling game outcomes of the Brazilian 2006 Series A championship as ordinal-valued. *Brazilian Journal of Probability Statistics*, v.22, n.2, p. 89-104, 2008.
- CLARKE, S. R.; NORMAN, J. M. Home ground advantage of individual clubs in english soccer. *The Statistician*, v.44, n.4, p. 509-521, 1995.
- DE FINETTI, B. *Probability, Induction and Statistics*. 1.ed. London: John Wiley, 1972. 300p.
- DYTE, D.; CLARKE, S. R. A ratings based Poisson model for world cup soccer simulation. *Journal of the Operational Research Society*, v.51, n.8, p. 993-998, 2000.
- GELMAN, A.; STURTZ, S.; LIGGES, U.; GORJANC, G.; KERMAN, J. *The R2winbugs Package Manual Version 2.0-4*. New York, Statistic Department Faculty, 2006.
- HOLGATE, P. Estimation for the bivariate Poisson distribution. *Biometrika*, v.51, p. 241-287, (1964).
- KARLIS, D.; NTZOUFRAS, I. Analysis of sports data by using bivariate Poisson models. *The Statistician*, v.52, n.3, p. 381-393, 2003.
- KARLIS, D.; NTZOUFRAS, I. Bayesian modelling of football outcomes: Using the Skellam's distribution for the goal difference. *IMA Journal of Management Mathematics*, v.20, n.2, p. 133-145, 2009.
- KELLER, J. B. A characterization of the Poisson distribution and the probability of winning a game. *The American Statistician*, v.48, n.4, p. 294-298, 1994.
- LEE, A. Modeling scores in the Premier League: Is Manchester United really the best?. *Chance*, v.10, n.1, p. 15-19, 1997.
- LOUZADA-NETO, F.; SUZUKI, A. K.; SALASAR, L. E. B. Predicting match outcomes in the English Premier League: Which will be the final rank?. *Journal of Data Science*, v.12, p. 235-254, 2014.
- POLLARD, R. Home advantage in soccer: A retrospective analysis. *Journal of Sports Sciences*, v.4, n.3, p. 237-248, 1986.
- R CORE TEAM. *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing. Vienna, Austria, 2016.
- SPIEGELHALTER, D. J.; THOMAS, A.; BEST, N. G. *WinBUGS Manual [version 1.4.1]*. MRC Biostatistics Unit, Cambridge University, UK, 2003.
- SUZUKI A. K.; SALASAR, L. E. B.; LOUZADA-NETO, F.; LEITE, J. G. A bayesian approach for predicting match outcomes: The 2006 (association) football world cup. *Journal of the Operational Research Society*, v.61, n.10, p.1530-1539, 2009.

Recebido em 15.12.2015.

Aprovado após revisão em 23.09.2016.

Apêndice A: Estimativas e Gráficos das autocorrelações

Neste apêndice apresentamos as estimativas para os parâmetros α_m e α_v para o jogo da vigésima rodada entre Manchester United (m) e West Bromwich (v) e para o jogo da trigésima rodada entre Manchester United (m) e Reading (v), para $t = \{m, v\}$ em que t^c representa o complementar de t .

As Tabela 9 e 10 mostram as estimativas dos parâmetros para o jogo da 20^a e 30^a rodada, respectivamente.

Tabela 9 - Estimativa (Desvio Padrão)

Time	Parâmetro			
	μ	α_c	α_{at}	α_{dt^c}
Manchester united	0,082 (0,068)	0,268 (0,088)	0,648 (0,143)	0,120 (0,187)
West Bromwich	0,082 (0,068)	–	0,077 (0,188)	-0,045 (0,208)

Tabela 10 - Estimativa (Desvio Padrão)

Time	Parâmetro			
	μ	α_c	α_{at}	α_{dt^c}
Manchester United	0,126 (0,054)	0,258 (0,071)	0,556 (0,125)	-0,167 (0,174)
Reading	0,126 (0,054)	–	-0,106 (0,168)	0,353 (0,132)

As Figuras 4 e 5 mostram os gráficos das autocorrelações estimadas. A Figura 4 mostra os gráficos das autocorrelações estimadas para os parâmetros associados ao número esperado de gols do jogo entre Manchester United e West Bromwich realizado na vigésima rodada do campeonato. A Figura 5 mostra o gráfico das autocorrelações estimadas para o jogo entre manchester United e Reading.

As Figuras 4(a,b) e 5(a,b) mostram as autocorrelações estimadas para o intercepto μ e para o fator casa α_c . As Figuras 4(c,d) e 5(c,d) mostram as autocorrelações estimadas para o efeito de ataque do time mandante e defesa do time visitante, $(\alpha_{am}, \alpha_{dm})$; e as Figuras 4(e,f) e 5(e,f) mostram o efeito de ataque do time visitante e defesa do time mandante, $(\alpha_{av}, \alpha_{dv})$, respectivamente. Note que, as autocorrelações são não significativas. Este fato também ocorre para os outros times em todas as outras rodadas previstas.

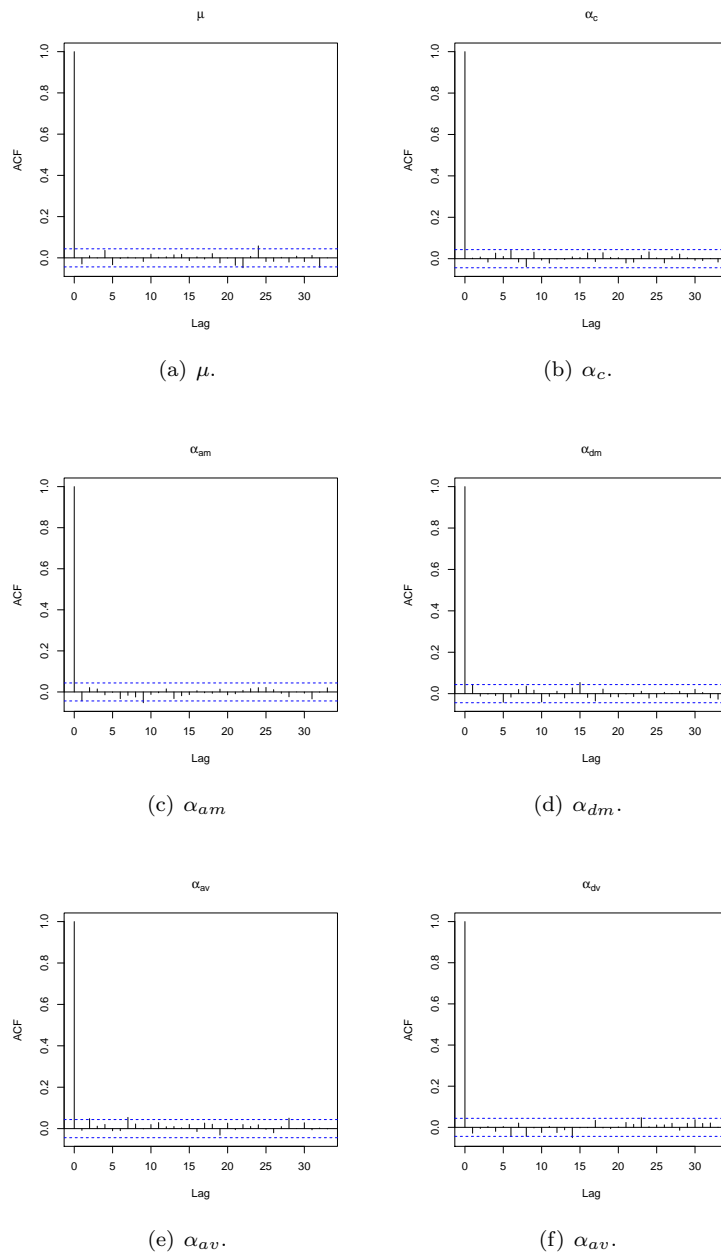


Figura 4 - Autocorrelações estimadas $(\alpha_{am}, \alpha_{dm})$ e $(\alpha_{av}, \alpha_{dv})$.

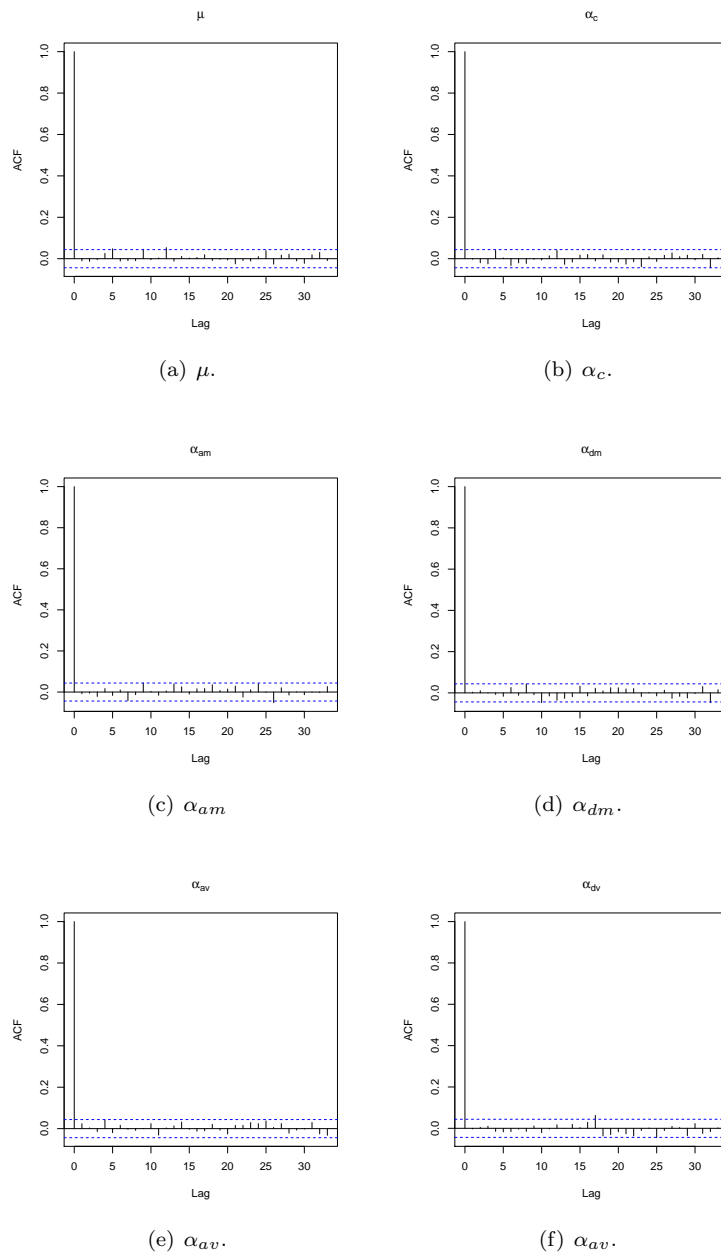


Figura 5 - Autocorrelações estimadas (α_{am}, α_{dm}) e (α_{av}, α_{dv}).

Apêndice B: Código WinBUGS

Neste apêndice apresentamos o código WinBUGS utilizado neste trabalho.

```
model{  
  
  for(i in 1:n){  
    x[i] ~ dpois(lambdaX[i])  
    y[i] ~ dpois(lambdaY[i])  
    log(lambdaX[i]) <- mu + casa + ata[man[i]] + def[vis[i]]  
    log(lambdaY[i]) <- mu + ata[vis[i]] + def[man[i]]  
  }  
  
  ata[1] <- - sum(ata[2:20])  
  def[1] <- - sum(def[2:20])  
  
# Distribuições a priori  
  
  mu ~ dnorm(0, 0.001)  
  casa ~ dnorm(0, 0.001)  
  
  for(i in 2:20){  
    ata[i] ~ dnorm(0, 0.001)  
    def[i] ~ dnorm(0, 0.001)  
  }  
  
}
```

Os times ordenados por ordem alfabética: Arsenal, Aston, Chelsea, Everton, Fulham, Liverpool, Manchester City, Manchester United, Newcastle, Norwich, Queens Park Rangers, Reading, Southampton, Stoke, Sunderland, Swansea, Tottenham, West Bromwich, West Ham, Wigan. Neste trabalho, consideramos os times enumerados de acordo com a sua posição em ordem alfabética. Por exemplo, o Arsenal é o time 1, o Chelsea é o time 3, enquanto que o West Ham é o time 19.

Os dados para todos os 380 jogos:

```
# Times mandantes
```

```
man=c(1,4,5,7,9,11,12,18,19,20,2,3,6,8,10,13,14,15,16,17,3,6,7,9,13,16,17,18,19,20,1,  
2,4,5,8,10,11,12,14,15,3,6,7,9,13,16,17,18,19,20,1,2,4,5,8,10,11,12,14,15,3,6,7,9,13,16,  
17,18,19,20,5,6,8,10,11,15,16,17,18,19,1,2,3,4,7,9,12,13,14,20,5,6,8,10,11,16,15,17,18,  
19,1,2,3,4,7,9,12,13,14,20,1,5,6,7,9,10,11,12,18,19,2,3,4,8,13,14,15,16,17,20,2,8,13,14,15,  
16,17,20,3,4,1,5,6,7,9,10,11,12,18,19,1,2,4,5,7,13,15,16,19,20,3,6,8,9,10,11,12,14,17,18,
```

3,6,7,9,13,16,17,18,19,20,1,2,4,5,8,10,11,12,14,15,1,2,4,5,8,10,11,12,14,15,3,6,7,9,13,16,
17,18,19,20,1,2,4,5,8,10,11,12,14,15,3,6,7,9,13,16,17,18,19,20,1,2,8,10,11,12,14,15,4,5,1,
4,5,7,9,11,12,18,19,20,2,3,6,8,10,13,14,15,16,17,1,5,6,7,9,10,11,12,18,19,2,3,4,8,13,14,15,
16,17,20,1,5,6,7,9,10,11,12,18,19,2,3,4,8,13,14,15,16,17,20,1,2,4,5,7,13,15,16,19,20,3,6,
8,9,10,11,12,14,17,18,1,2,3,4,7,9,12,13,14,20,5,6,8,10,11,15,16,17,18,19,1,2,3,4,7,9,12,13,
14,20,5,6,8,10,11,15,16,17,18,19,1,2,4,5,8,10,11,12,14,15,3,6,7,9,13,16,17,18,19,20)

Times visitantes

vis=c(15,8,10,13,17,16,14,6,2,3,4,9,7,5,11,20,1,12,19,18,12,1,11,2,8,15,10,4,5,14,13,16,9,
18,20,19,3,17,7,6,14,8,1,10,2,4,11,12,15,5,3,18,13,7,17,6,19,9,16,20,10,14,15,8,5,12,2,11,
1,4,2,12,14,1,4,9,20,3,7,13,11,10,8,6,16,18,5,17,15,19,4,9,1,14,12,3,2,20,13,7,5,8,6,15,17,
19,10,16,11,18,17,15,20,2,16,8,13,4,3,14,1,7,10,11,9,5,18,6,19,12,12,19,10,9,11,18,6,7,5,
1,16,17,13,4,20,15,2,8,14,3,18,14,17,9,8,12,3,10,6,11,13,2,15,7,20,5,1,4,16,19,2,5,12,11,
15,8,14,10,4,1,19,17,20,13,9,3,18,16,6,7,9,20,3,16,18,7,6,19,13,17,11,15,14,4,1,2,12,5,10,
8,7,13,16,20,6,9,17,18,3,19,1,10,5,12,4,14,8,2,11,15,6,9,13,17,7,3,20,16,18,19,14,2,8,6,3,
10,15,17,16,13,19,20,18,4,5,7,12,1,11,9,2,14,16,3,13,4,8,20,15,17,7,18,12,10,11,19,5,9,1,
6,4,3,17,20,14,13,15,2,16,8,11,19,7,12,6,18,10,1,5,9,12,6,14,11,9,3,8,17,18,10,15,19,7,5,
16,20,13,2,4,1,10,5,17,11,18,15,6,19,8,16,1,3,2,12,14,4,13,7,9,20,8,15,16,5,19,6,11,18,10,
17,12,4,3,2,1,14,7,13,20,9,20,3,19,6,16,18,9,7,17,13,4,11,10,1,14,5,15,8,12,2)

Gols marcados pelos times mandantes

x=c(0,1,5,3,2,0,1,3,1,0,1,2,2,3,1,0,0,3,3,1,4,0,3,1,2,2,1,2,3,2,6,2,2,3,4,0,0,1,1,1,1,1,1,4,
0,2,1,1,1,1,3,1,2,2,1,2,2,1,4,0,3,0,2,2,2,3,1,2,1,1,4,1,1,1,2,2,1,4,1,1,2,2,1,2,3,1,0,2,2,1,2,
1,1,1,0,0,2,0,3,2,1,2,2,0,0,1,1,1,5,1,3,5,1,1,1,2,2,1,0,0,1,3,2,1,2,0,3,3,1,1,1,2,0,3,2,0,0,1,0,
0,1,1,3,2,1,3,0,3,2,0,2,2,2,1,1,3,2,2,2,1,3,1,2,2,2,1,1,0,8,4,1,1,0,1,0,2,1,0,5,0,2,1,4,0,1,0,3,
1,7,0,1,1,2,3,0,1,3,1,0,3,3,1,1,2,3,1,2,0,0,0,0,1,2,0,0,3,0,3,2,5,2,1,0,3,1,2,1,2,2,1,2,1,0,2,2,
0,2,3,1,3,0,2,3,0,2,0,1,2,2,4,0,2,0,3,2,0,4,2,2,1,5,2,4,2,0,0,2,2,0,1,3,4,1,0,2,1,2,0,0,0,3,1,2,
0,3,1,2,2,3,2,2,1,3,0,1,0,0,2,4,1,1,3,4,2,0,1,3,1,2,0,1,1,2,1,0,1,2,1,3,1,2,2,1,0,0,1,0,2,0,2,3,
2,0,1,0,3,1,2,1,6,2,1,2,0,0,0,1,2,2,0,0,1,0,1,0,1,2,0,4,1,2,1,2,4,1,0,1,1,2,1,2,0,1,0,1,5,4,2)

Gols marcados pelos times visitantes

y=c(0,0,0,2,1,5,1,0,0,2,3,0,2,2,1,2,0,0,0,1,2,2,1,1,3,2,1,0,0,2,1,0,2,0,0,0,0,3,1,1,0,2,1,0,1,
3,1,0,1,2,2,1,1,2,3,5,2,2,0,0,1,0,0,3,2,2,0,2,3,2,0,0,2,0,1,1,1,4,2,1,0,1,3,2,0,1,3,2,0,1,2,1,1,
0,1,1,1,1,0,0,3,3,1,1,1,1,0,1,0,2,2,3,0,0,2,0,3,1,1,1,0,0,1,1,0,0,4,0,1,2,0,0,1,1,0,1,1,2,0,1,2,
3,0,1,0,1,1,4,1,1,0,0,1,1,3,0,3,4,3,2,2,3,1,3,1,1,5,1,0,0,0,0,0,0,1,1,0,1,2,1,1,4,1,1,3,1,2,0,1,
0,3,3,2,2,0,4,3,0,3,2,1,0,0,2,1,2,1,2,1,4,2,1,0,1,1,0,0,2,4,0,1,0,0,2,0,1,1,2,1,3,2,2,1,1,0,2,2,
0,1,1,0,3,1,2,2,0,1,1,0,2,1,1,2,0,0,1,1,1,1,1,0,0,0,2,1,2,3,1,3,1,0,1,0,2,1,2,0,1,4,0,3,2,0,1,
0,1,2,1,2,2,0,0,0,1,0,1,2,1,1,1,2,0,2,0,1,1,2,1,0,1,0,2,0,2,1,2,3,2,2,1,1,2,0,0,3,0,1,2,3,1,2,0,
1,2,0,0,1,1,0,1,1,0,0,1,6,0,3,0,2,4,0,1,2,1,1,0,0,3,0,1,2,0,3,1,0,2,2,1,1,1,0,3,1,1,3,0,5,2,2)