

MS777 - PROJETO SUPERVISIONADO

# TÉCNICA PARA PREDIÇÃO DE PARTIDAS DE FUTEBOL



**Aluno:** Carlos Danilo Tomé - RA 145640  
**Orientador:** Prof. Dr. Laércio Luis Vendite

Universidade Estadual de Campinas - UNICAMP  
2º semestre de 2018

## ***1 Introdução***

A origem de esportes similares ao futebol é desconhecida, mas se pode citar relatos sobre a prática esportiva que consistia em usar os pés no controle da bola com o objetivo de acertar uma baliza em diversas culturas e regiões pelo mundo, e o registro documentado mais antigo da prática desse esporte data dos séculos segundo e terceiro A.C. na China durante a dinastia Han e se chamava Tsu ‘Chu, e tinha sua prática ligada a exercícios militares. Outros antecessores similares ao futebol moderno também foram praticados na Grécia, Japão e no Império Romano. Durante o século XIV já se praticava similares ao futebol em diversos países da Europa mas com regras distintas, e foi na Inglaterra em 1863 que se criou a “*Football Association*” que padronizou as regras e práticas do futebol como esporte [1], um importante marco na história desse esporte pois criou condições para disseminação e popularização mundial.

Com o passar dos séculos o Futebol se tornou um dos esportes mais praticados pelo mundo e um valioso produto, despertando interesses dos mais variados. Segundo [2] relatório divulgado pela FIFA (*Fédération Internationale de Football Association*) foi movimentado em transferências de atletas cerca de 6.3 US\$ bilhões somente durante o ano de 2017, a Copa do Mundo Rússia 2018 obteve cerca de 3,2 bilhões de espectadores ao longo do torneio e é considerado o segundo maior evento esportivo do planeta, atrás somente das Olimpíadas de Verão em quantidade de espectadores.

Na literatura existe uma vasta gama de trabalhos voltados à análise e desenvolvimento do futebol, um dos primeiros trabalhos que sugere modelos de predição para determinar resultados de partidas de futebol foi escrito por Maher em 1982 [3], introduzindo na literatura a idéia de definir a quantidade de gols do time  $i$  na partida contra o time  $j$  como  $X_{ij}$  e similarmente  $Y_{ij}$  como a quantidade de gols do seu adversário na mesma partida, e assumir  $X_{ij}$  e  $Y_{ij}$  variáveis independentes. Outra importante ideia de Maher foi definir a força de ataque do time  $i$  como  $\alpha_i$  e  $\beta_i$  como força de defesa do time  $i$ , e por fim,  $\gamma$  como fator mando de campo. O modelo que descreve a probabilidade do resultado da partida do Time mandante  $i$  contra o Time  $j$ :

$$P(X_{ij} = x, Y_{ij} = y | \alpha, \beta, \gamma) = Poisson(x | \gamma \cdot \alpha_i \beta_j) \cdot Poisson(y | \alpha_j \beta_i)$$

No qual a notação  $Poisson(z|\lambda)$  serve para descrever a função distribuição de probabilidade de Poisson com parâmetro  $\lambda$  para avaliar  $z$ . Esta sugestão de descrever a probabilidade dos resultados por meio da força de ataque/defesa e fator mando de campo é amplamente utilizado na literatura.

Em [4] Langseth propõe uma estratégia para realizar apostas esportivas utilizando predição de resultados sugerido por Dixon e Coles [5], este por sua vez sugere que a predição de resultados esportivos pode ser feita por meio de um modelo com dependência temporal, usando dados de habilidade de ataque e defesa dos times para determinar o parâmetro  $\lambda$  da distribuição de Poisson que descreve a quantidade de gols em uma partida, atualizando os parâmetros a cada rodada.

Existem ainda na literatura outras abordagens para modelar predição de partidas de futebol utilizando de técnicas como método de Monte Carlo baseado em cadeias de Markov e inferência Bayesiana como em [17] ou métodos baseados em Data Mining e Machine Learning como abordagem para predição de partidas de futebol como utilizada em Bezerra, de Abreu e Esmín em [10] e Duarte em [11].

Para este projeto foi utilizado o método Soma e Diferença [SD0] proposto inicialmente por Arruda em [6], e reaplicado ao Campeonato Brasileiro de 2007 por Suzuki e Tavares em [14], adicionando o fator crise no método. Este método propõe a idéia de assumir que o número de gols marcados por um time em uma partida obedece a uma distribuição univariada de Poisson, utilizando dados históricos para determinar o parâmetro desta distribuição, são eles: média de gols do campeonato, poder ofensivo da equipe, poder defensivo do adversário e fator mando de campo. Para descrever o resultado da partida o método propõe a utilização da distribuição de Poisson Bivariada de Holgate.

## 2 Modelo

### 2.1 Modelo para Distribuição de Poisson

Neste projeto será adotado a suposição de que a quantidade de gols marcados pela equipe mandante denotado por  $X$  e pela equipe visitante por  $Y$  possua alguma distribuição discreta bivariada de Poisson, portanto para o vetor de probabilidade  $(X, Y)$

$$\sum_{x=0}^{\infty} \sum_{y=0}^{\infty} P(X = x, Y = y) = 1$$

$$\sum_{x=0}^{\infty} P(X = x, Y = y) = \frac{e^{-\lambda_y} (\lambda_y)^y}{y!}$$

$$\sum_{y=0}^{\infty} P(X = x, Y = y) = \frac{e^{-\lambda_x} (\lambda_x)^x}{x!}$$

com  $\lambda_x, \lambda_y > 0$

Segundo modelo proposto por Arruda em [6] a classe das distribuições de Poisson que melhor se adapta à aplicação ao futebol é a distribuição “de Holgate”. Uma observação importante à se fazer segue do fato que a covariância é nula, implicando que  $X$  e  $Y$  são eventos independentes. Definimos a esperança marginal de gols a serem marcados em cada jogo como

$$E[X] = \lambda_1 + \lambda_{12}$$

$$E[Y] = \lambda_2 + \lambda_{12}$$

Mas,

$$\begin{aligned} \text{Cov}(X, Y) = 0 &\Rightarrow \lambda_{12} = 0 \\ &\Rightarrow X \text{ e } Y \text{ independentes.} \end{aligned}$$

O próximo passo neste método é estimar os parâmetros  $\lambda_1$  e  $\lambda_2$ , através do passo que denomina este método, ou seja, Soma e Diferença das esperanças de X e Y. Como visto no passo anterior  $E[X] = \lambda_1$  e  $E[Y] = \lambda_2$  e portanto, é possível escrever a seguinte expressão:

$$E[X - Y] = \lambda_1 - \lambda_2 \quad (1a)$$

$$E[X + Y] = \lambda_1 + \lambda_2 \quad (1b)$$

Sendo a expressão (1a) relacionado ao quanto é esperado a diferença de gols entre Mandante e Visitante e similarmente, (1b) relacionado a soma de gols esperado em uma partida do Campeonato. Para determinar os estimadores indiretos, que é objeto do interesse deste método ao qual é parâmetro para a distribuição de Poisson, sugere-se um modelo linear dado que é composto por vetores de incógnitas relacionado à qualidade ofensiva e defensiva dos time e bem como o fator mandante.

## 2.2 Método SD 0

Arruda em [6] definiu quatro métodos para previsão de resultados de partida de futebol dois deles implícitos, Método SD 0 e Método SD 1, e os outros dois métodos explícitos, Método Chance I e Método Chance II. Na literatura é comum observar que autores têm utilizado dos métodos implícitos para realizar previsões, seja pela maior facilidade prática na manipulação e determinação dos vetores de incógnitas ou pela descrição simples e poderosa que este modelo descreve neste tipo de problema.

**Considere o Modelo SD 0:**

$$(X + Y)_j = S_j \mu + \gamma + erro_{Sj} \quad (2a)$$

$$(X - Y)_j = D_j \phi + \gamma + erro_{Dj} \quad (2b)$$

Com,

$(X + Y)_j =$  Soma de gols marcados na partida j.  
 $j = 1, 2, 3, \dots, \#$  de jogos no banco de dados.

$T = 1, 2, 3, \dots$ , # de times participantes da competição

$S_j =$  Matriz linha status do jogo para a Soma de gols.

$D_j =$  Matriz linha status do jogo para a Diferença de gols.

$erro_{D_j}$ ,  $erro_{S_j} =$  Erros independentes; com média zero.

$\mu$ ,  $\phi =$  Vetor de parâmetros.

$\gamma =$  Fator mando de campo

Utilizando o banco de dados construí-se os sistemas  $(X + Y)_j$  e  $(X - Y)_j$ , em que  $S_j$  é a matriz com T colunas e j linhas, com valores correspondente à 1 se o clube representado pela coluna irá participar da partida e 0 caso contrário. E a matriz similar  $D_j$  que atribui valores de -1 caso o clube correspondente a aquela coluna seja visitante, 1 caso o clube correspondente a aquela coluna for mandante e 0 caso contrário.

O vetor de mando de campo  $\gamma$  atribui valor 1 se o mando de campo for exercido efetivamente pelo clube mandante daquela partida e 0 caso contrário (como jogo de portões fechados, partida em campo neutro, e etc). A adição do fator mando de campo, que matematicamente representa maior chance de  $(X + Y)_j$  e  $(X - Y)_j$  serem positivos implicando no acréscimo em  $E[X]$  como também na probabilidade do Time Mandante vencer a partida.

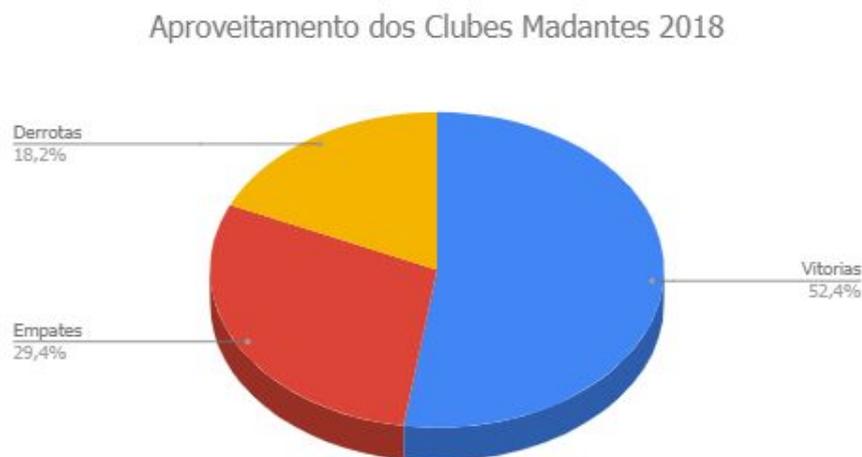


Gráfico 1: Aproveitamento dos Times Mandantes no Campeonato Brasileiro de 2018 até a 34ª rodada.

Como podemos analisar no Gráfico 1 times que exercem o mando de campo tendem a ter maiores chances de ganhar a partida, fato embasado no senso comum de torcedores e especialistas em que o clube mandante tem chances maiores de vencer o jogo, intuição que foi investigada em [9] no qual o autor verifica a vantagem para clubes que mandam a partida em sua casa através da porcentagem de vitórias em dados históricos e ainda observou que essa vantagem diminui em divisões inferiores de um mesmo campeonato.

Por fim, o vetor  $\mu$  construído igualmente ao vetor  $\varphi$ , com T linhas correspondente ao fator qualidade  $\mu$  do clube representado por aquela linha.

Resumindo o problema de encontrar  $\lambda_1$  e  $\lambda_2$  em determinar os vetores parâmetros  $\mu$  e  $\varphi$ , para a estratégia utilizada é minimizar  $erro_{Sj}$  e  $erro_{Dj}$  através do método dos quadrados mínimos adequado para este problema. Pois em posse dos parâmetros  $\mu$  e  $\varphi$  fixados pelo banco de dados podemos prever a esperança em partidas futuras  $E'[X - Y]$  e  $E'[X + Y]$  com passo final determinar a matriz de probabilidade “de Holgate”, definido em [16], com os placares prováveis.

### ***3 Desenvolvimento***

O Campeonato Brasileiro é disputado anualmente por 20 clubes no formato pontos corridos na qual cada time enfrenta todos os outros participantes do torneio em dois turnos, como Mandante (sediando a partida em seu estádio) e Visitante, atribuindo 3 pontos pro vencedor da partida, 1 ponto para ambos caso tenham empatado e 0 pontos para o clube perdedor, o clube que somar mais pontos ao final das 38 rodadas do torneio sagra-se campeão do campeonato.

#### ***3.1 Exemplo para estimação de $\lambda_1$ e $\lambda_2$***

Utilizando como exemplo alguns jogos da décima rodada do Brasileirão 1, com jogos realizados entre os dias 5 à 7 de junho,

### ***Torneio Hipotético***

Sport 1 x 0 Atlético Paranaense	$(X + Y)_{94} = 1$	$(X - Y)_{94} = 1$
Corinthians 1 x 1 Santos	$(X + Y)_{95} = 2$	$(X - Y)_{95} = 0$
Grêmio 0 x 2 Palmeiras	$(X + Y)_{97} = 2$	$(X - Y)_{97} = -2$
América 1 x 2 Atlético Mineiro	$(X + Y)_{100} = 3$	$(X - Y)_{100} = -1$

Neste modelo os parâmetros  $\mu$  e  $\varphi$  tem que ser calibrados separadamente, sendo assim se utilizarmos como banco de dados um campeonato curto somente com  $j = 4$  o problema se resumiria à minimizar  $erro_{D_j}$  e  $erro_{S_j}$  nas funções (2a) e (2b). Nas quais o vetores seriam,

$$(X+Y)_j = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix} \quad (X-Y)_j = \begin{bmatrix} 1 \\ 0 \\ -2 \\ -1 \end{bmatrix}$$

$$S_j = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$D_j = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Note que na última coluna foi adicionado o parâmetro fator mando de campo, segue a expressão para o vetor de incógnitas

$$\mu^T = [\mu_{AME}, \mu_{ATR}, \mu_{ATM}, \mu_{COR}, \mu_{GRE}, \mu_{PAL}, \mu_{SAN}, \mu_{SPO}, \gamma_{MANDO DE CAMPO 1}]$$

$$\varphi^T = [\varphi_{AME}, \varphi_{ATR}, \varphi_{ATM}, \varphi_{COR}, \varphi_{GRE}, \varphi_{PAL}, \varphi_{SAN}, \varphi_{SPO}, \gamma_{MANDO DE CAMPO 2}]$$

Note que o módulo do fator mando de campo é comum para todos os clubes, mas difere nos sistemas para a 2.a e 2.b .

Conforme definido por Monroe em 1920 [12], existe uma matriz inversa generalizada  $A$  e provou sua unicidade. Da teoria de matriz inversa generalizadas de Monroe-Penrose, Zontini em [13] propôs métodos computacionais de a fim de determinar a inversa generalizada de Monroe-Penrose  $A$ , portanto do problema de minimização do  $erro_{Dj}$  e  $erro_{Sj}$  , podemos escrever que os vetores deste exemplo como

$$\mu = (S^*S)^{-1}S(X+Y) = A S'(X+Y) \Rightarrow \mu^T = [1.25, 0.25, 1.25, 0.75, 0.75, 0.75, 0.75, 0.25, 0.5]$$

Onde a inversa generalizada de  $(S^*S)$  é  $A$ . Similarmente para  $D$  ,

$$\varphi = (D'D)^{-1}D(X-Y) = B D'(X-Y) \Rightarrow \varphi^T = [0.25, 0.25, 1.5, 0.5, 0.25, 0.25, 0.75, 1.0, 0.25]$$

Logo, podemos obter os estimadores indiretos

$$\lambda_1' = \frac{E'[X+Y] + E'[X-Y]}{2}$$

$$\lambda_2' = \frac{E'[X+Y] - E'[X-Y]}{2}$$

Outra consequência importante que é vital para este método aborda a invariância dos parâmetros atribuído ao vetor  $\mu$  e  $\varphi$  segundo Monroe em [12] pois seja  $M\bar{v} = y$ , com múltiplas soluções , qualquer escolha que atenda a fórmula  $\bar{v} = (M^*M)^{-1}M$  é verdadeira para  $m\bar{v}$ , onde  $m$  pertence ao espaço gerado por  $M$ .

### **3.2 Predição de resultados**

Para prever a quantidade de gols do time mandante  $X$  e a quantidade de gols marcados pelo visitante  $Y$  em um jogo hipotético entre Corinthians e Palmeiras, com mando de campo para a equipe do Corinthians

$$E' [X] = 0.125$$

$$E' [Y] = 3.75$$

Obviamente é impossível observar um resultado não discreto em partidas de futebol, mas esse é o placar dos “quadrados mínimos” ainda se faz necessário construir as probabilidades discretas já que foi determinado suficientemente  $\lambda_1$  e  $\lambda_2$  .

Ademais pode se calcular a probabilidade para eventuais jogos a serem realizados através da propriedade de probabilidade conjunta, utilizando a distribuição de probabilidade de Holgate. Na prática trunca-se  $x,y$  em 10, pois a incidência de partidas em que são marcados 10 gols é ínfima. Segue que,

(i) *Probabilidade do time mandante vencer a partida*

$$\sum_{x=0}^{\infty} \sum_{y=0}^{\infty} P(x > y)$$

(ii) *Probabilidade do time visitante vencer a partida*

$$\sum_{x=0}^{\infty} \sum_{y=0}^{\infty} P(x < y)$$

(iii) *Probabilidade de empate*

$$\sum_{x=0}^{\infty} \sum_{y=0}^{\infty} P(x = y)$$

### **3.2 Qualidade das previsões**

Como sugerido por Arruda em [6] e baseado em sua ampla utilização na literatura para qualificar as previsões realizadas em eventos tricotômicos é razoável utilizar a medida da distância DeFinetti, enunciado em 1972 por Bruno DeFinetti [15] esta medida é construída acerca da representação geométrica de um simplex, ou seja um tetraedro regular [Figura1], e do ponto com

coordenadas definidas a partir das probabilidades de vitória do time mandante (PV), probabilidade de empate (PE) e da probabilidade de derrota do time mandante (PD), todas não nulas.

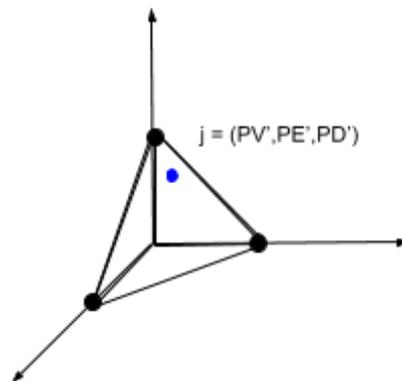


Figura 1: Tetraedro Regular e representação do ponto  $j = (PV', PE', PD')$ .

Os vértices desse simplex correspondem à ocorrência do evento, no caso do futebol o resultado observado da partida, e todos os demais pontos correspondem a previsões possíveis do evento. Definindo a distância de DeFinetti como a distância euclidiana quadrática para o vértice representante do resultado observado como o descrito abaixo,

$$TR = \{(PV, PE, PD) \in \mathbb{R}^3 : PV + PE + PD = 1\}$$

com  $PV > 0, PE > 0, PD > 0$

E portanto, a distância DeFinetti é a distância euclidiana quadrática do ponto  $j = (PV', PE', PD')$ , gerado pelas previsões de resultado do time mandante via Método SD 0, aos respectivos vértices

(0,0,1) caso o time mandante seja derrotado.

(0,1,0) caso o jogo termine empatado.

(1,0,0) caso o time mandante seja vitorioso na partida.

Note que se forem atribuídos probabilidades idênticas aos 3 resultados, ou seja  $(PV = PE = PD = \frac{1}{3})$ , a medida da distância de DeFinetti aos pontos equiprováveis é  $r = 0,6667$ . Este número é usado como base de comparação para a qualidade das previsões, se a medida de DeFinetti  $r' < r$  então as previsões que geraram  $r'$  são boas, caso contrário serão consideradas previsões de má qualidade.

Paralelamente é razoável qualificar as previsões segundo a taxa de acerto, comparando o evento observado com o evento previsto. Considera evento previsto como Vitória do mandante caso  $PV = \max\{PV, PE, PD\}$ , Empate caso  $PE = \max\{PV, PE, PD\}$  e derrota caso  $PD = \max\{PV, PE, PD\}$ .

### 3.3 Aplicação Campeonato Brasileiro 2018

Para formar o banco de dados referente aos jogos do Campeonato Brasileiro de 2018 foram analisados os  $j = 190$  primeiros jogos do nacional, ou seja, o primeiro turno completo da competição onde todos os clubes participantes já se enfrentaram, através de resultados extraídos do portal oficial da *Confederação Brasileira de Futebol* [7] na internet. Obtidos os vetores de parâmetros  $\varphi$  e  $\mu$  para os 20 times em ordem alfabética e por fim o fator mando de campo,

$$\mu^T = [1.07, 2.18, 0.95, 1.23, 1.12, 0.57, 1.12, 0.34, 1.18, 1.12, 0.45, 0.93, 1.18, 0.65, 1.18, 0.65, 1.18, 1.40, 1.29, 1.45, 2.29]$$

$$\varphi^T = [-1.03, -1.33, -0.1, -0.68, -1.45, -1.08, -1.25, -1.38, -0.55, -1.33, -1.25, -0.68, -1.08, -0.7, -1.2, -0.63, -1.0, -1.1, -0.68, -1.8, 2.28]$$

Foi simulado 30 jogos do campeonato obtém-se a seguinte tabela, considerando a Medida de De Finetti média dos jogos simulados, e considerando acerto quando o resultado observado bate com a maior probabilidade entre PV, PE e PD:

	Taxa de Acerto	Medida de DeFinetti
Método SD 0	58,33%	0,6422

Tabela 1: Taxa de Acerto nos 30 jogos simulados e Medida de DeFinetti média.

## ***4 Conclusão***

Portanto o Método SD 0, assim como Método SD 1, é uma ferramenta amplamente utilizada na literatura para previsão resultados em um esporte que tem como máxima sua imprevisibilidade e os diversos fatores que realmente influenciam no placar final. Arruda descreve de forma simples, porém poderosa, os fatores que influenciam no placar do jogo como fator mando de campo  $\gamma$ , poder ofensivo  $\mu$  e poder defensivo  $\varphi$ .

Analisando a Medida da Distância de DeFinetti observa-se que  $r' = 0,6422$  está muito próximo ao  $r = 0,6667$ , correspondente a distância DeFinetti equiprovável, mas ainda assim é menor o que garante uma boa qualidade nas previsões segundo o que foi proposto ao método. Todavia, a taxa de acerto das previsões não foi tão boa quanto o esperado, e neste aspecto o baixo número de jogos simulados e também a quantidade de jogos no banco de dados podem ter influenciado negativamente este resultado. Outras peculiaridades do campeonato brasileiro deste ano também influenciaram nos resultados como a mudança no calendário da Copa do Brasil e Libertadores que passaram a ser disputadas no ano inteiro, dividindo a atenção dos times com maior qualidade técnica.

Sendo assim a evolução natural deste projeto seria aumentar o banco de dados histórico considerando  $n$  campeonatos anteriores, adicionar o fator crise como fez Suzuki em [14] ou considerar um modelo dinâmico que atualizasse o vetor de parâmetros a cada rodada como proposto por Gamerman e Junior em [17] utilizando MCMC para determinar o vetor de incógnitas.

## Referências

- [1] <https://www.fifa.com/about-fifa/who-we-are/the-game/index.html> - acessado dia 12/11/2018.
- [2] [https://www.fifatms.com/wp-content/uploads/dlm\\_uploads/2018/01/GTM\\_2018.pdf](https://www.fifatms.com/wp-content/uploads/dlm_uploads/2018/01/GTM_2018.pdf) - acessado dia 12/11/2018.
- [3] MAHER, M. J. Modelling association football scores. **Statistica Neerlandica**, 36(3): p.109-118, 1982.
- [4] LANGSETH, H. Beating the bookie: A look at statistical models for prediction of football matches. **Frontiers in Artificial Intelligence and Applications**. 257. p 165-174, 2013.
- [5] DIXON, M. J. ; COLES, S. G. **Modeling association football scores and inefficiencies in the football betting market**. *Applied Statistics*, 46:265–280, 1997.

- [6] ARRUDA, L. **Poisson, Bayes, Futebol e DeFinetti**. Dissertação (Mestrado em Estatística), Instituto de Matemática e Estatística da Universidade de São Paulo, 2000.
- [7] <https://www.cbf.com.br/> - acessado dia 12/11/2018.
- [8] Bååth, R. **Modeling Match Results in Soccer using a Hierarchical Bayesian Poisson Model**. LUCS Minor, 18. (2015) (ISSN 1104-1609)
- [9] POLLARD, R.: **Home advantage in soccer: a retrospective analysis**. Journal of Sports Sciences, pp. 237 – 248, 1986.
- [10] BEZERRA S,C. A. C. ; de Abreu, A. A. A. M. ; Esmín, A. A. A. **Previsão de Resultados de Jogos do Campeonato Brasileiro de Futebol: Uma Abordagem de Mineração de Dados**. XII Encontro Nacional de Inteligência Artificial e Computacional.(2015)
- [11] DUARTE, L. M. da S. **1X2 Previsão de Resultados de Jogos de Futebol**. International Conference on Machine Learning Applications. (2015)
- [12] MOORE, E. H. **On the reciprocal of the general algebraic matrix**, Bull, Amer. Math. Soc., 26 (1920), 394-395.
- [13] ZONTINI, D. E. **Métodos Computacionais para inversas generalizadas**. (2014)
- [14] SUZUKI, A. K. ; TAVARES, L. **Modelagem estatística para previsão esportiva: Uma aplicação no Futebol**. (2015) 32-47
- [15] DeFINETTI, B. (1972), **Probability, Induction and Statistics**, London: John Wiley.
- [16] HOLGATE, P. (1964), Estimation for the Bivariate Poisson Distribution, Biometrika 51, 241-245.
- [17] GAMERMAN, D. ; JUNIOR, O. G. de S. **Previsão de partidas de futebol usando modelos dinâmicos**. XXXVI Sociedade Brasileira de Pesquisa Operacional. (2014)