

# MÉTODO DE MAHER E O CAMPEONATO BRASILEIRO

Giancarlo A. R. Tosto    Rafael F. Spaziani    Thiago de  
Souza Duarte

Tratamento Estatístico de Dados em Física Experimental  
Prof. Zwinglio  
2016

# Visão Geral

- 1 Análise estatística de ligas e torneios de futebol
- 2 O Método de Maher
  - Hipóteses e Detalhes Teóricos
  - Implementação
  - Resultados
- 3 Conclusões
  - Conclusões

# Introdução

- Estudo matemático de jogos de aposta está relacionado ao **surgimento da probabilidade**.
- Historicamente há grande **interesse por apostas e jogos de azar** na Inglaterra
  - ~ **1660**: registros das **primeiras casas de apostas** na Inglaterra.
  - ~ **1850**: a "Era do *pedestrianism*"  $\Rightarrow$  apostas, instituições de atletismo.
  - ~ **1880**: **futebol** passa por processo de reformulação e profissionalização na Inglaterra
  - ~ **1950**: M. J. Moroney sugere relação entre distribuição de gols de partidas de futebol e uma **distribuição Poisson** modificada.
  - o **1982**: **M. J. Maher** publica o trabalho com seu modelo.

- Moroney e Maher eram ingleses e em particular o modelo do Maher usa dados do campeonato inglês por isso decidimos verificar se a hipótese da Poisson também vale para outros campeonatos, em especial a Série A do Campeonato Brasileiro.
- O Professor Zwinglio forneceu os dados que usamos para confirmar essa hipótese. Mas antes investigamos algumas características mais gerais de ligas e torneios entre 2000 e 2006 na maioria dos casos

⋮

# Competições

ALE	BRA-A	BRA-B	CAR	COP-B	ESP	FRA	ING	ITA	PAR	PAU	POR
1997	1997				1997	1997	1997	1997			1997
	1998										
	1999		1999	1999							
2000	2000			2000	2000	2000	2000	2000			2000
2001	2001		2001			2001	2001		2001	2001	2001
2002	2002		2002	2002		2002			2002		2002
2003	2003			2003	2003	2003	2003	2003			2003
	2004		2004		2004	2004	2004	2004	2004		2004
2005	2005	2005	2005	2005	2005	2005	2005	2005	2005	2005	2005
2006	2006	2006	2006	2006		2006	2006		2006	2006	2006
	2007	2007	2007	2007					2007	2007	
	2008	2008							2008		
	2009	2009									
	2010										
	2011										
	2012										
	2013										
	2014										
	2015										

Tabela 1: Dados usados (sem copa do mundo e com campeonato inglês incompleto).

# Competições

## Gols

Campeonatos	Número médio de gols	Saldo médio
Alemão	2.84 (0.11)	0.44 (0.16)
Brasileiro A	2.73 (0.22)	0.54 (0.09)
Brasileiro B	2.83 (0.04)	0.70 (0.08)
Carioca	3.09 (0.27)	0.42 (0.14)
Copa do Brasil	3.03 (0.13)	0.58 (0.19)
Francês	2.28 (0.13)	0.48 (0.10)
Espanhol	2.64 (0.16)	0.43 (0.16)
Inglês	2.58 (0.09)	0.41 (0.07)
Italiano	2.67 (0.11)	0.40 (0.07)
Paranaense	3.10 (0.26)	0.56 (0.13)
Paulista	3.23 (0.27)	0.40 (0.09)
Português	2.44 (0.18)	0.44 (0.10)

Tabela 2: Número médio de gols e saldo médio por partida

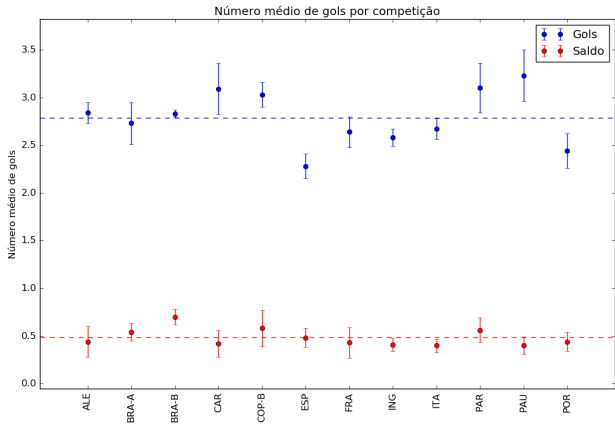


Figura 1: Número médio de gols e saldo médio por partida.

# Competições

## Resultados

Campeonatos	Mandante	Empate	Visitante
Alemão	47.2 (1.1) %	25.6 (0.9) %	27.2 (1.0) %
Brasileiro A	50.8 (0.6) %	25.6 (0.5) %	23.6 (0.5) %
Brasileiro B	55.7 (1.2) %	23.3 (1.0) %	21.0 (1.0) %
Carioca	49.8 (2.0) %	22.3 (1.7) %	27.9 (1.8) %
Copa do Brasil	49.0 (1.7) %	24.2 (1.5) %	26.8 (1.5) %
Francês	48.1 (0.9) %	29.2 (0.9) %	22.7 (0.8) %
Espanhol	47.6 (1.1) %	26.9 (1.0) %	25.4 (1.0) %
Inglês	46.8 (1.0) %	25.8 (0.9) %	27.3 (0.9) %
Italiano	45.1 (1.2) %	29.8 (1.1) %	25.0 (1.1) %
Paranaense	50.2 (1.7) %	25.5 (1.5) %	24.4 (1.5) %
Paulista	48.2 (1.7) %	23.0 (1.4) %	28.9 (1.5) %
Português	48.6 (1.0) %	26.0 (0.9) %	25.4 (0.9) %

Tabela 3: Frequências relativas médias de resultados por competição.



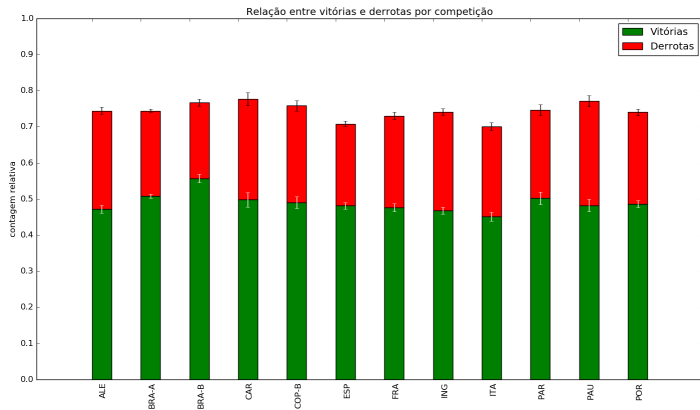


Figura 2: Relação entre vitórias e derrotas para diferentes competições.

# Campeonato Brasileiro A

## Gols

Brasileiro	Número de gols	Saldo médio
1997	2.76 (0.10)	0.62 (0.08)
1998	2.87 (0.10)	0.45 (0.09)
1999	2.84 (0.11)	0.38 (0.11)
2000	2.92 (0.09)	0.46 (0.09)
2001	2.86 (0.09)	0.48 (0.09)
2002	3.02 (0.10)	0.59 (0.09)
2003	2.88 (0.08)	0.67 (0.07)
2004	2.78 (0.07)	0.65 (0.07)
2005	3.13 (0.09)	0.47 (0.08)
2006	2.71 (0.09)	0.47 (0.08)
2007	2.76 (0.08)	0.58 (0.09)
2008	2.72 (0.09)	0.74 (0.09)
2009	2.88 (0.09)	0.59 (0.08)
2010	2.57 (0.08)	0.48 (0.08)
2011	2.68 (0.08)	0.53 (0.08)
2012	2.47 (0.08)	0.47 (0.08)
2013	2.46 (0.08)	0.47 (0.08)
2014	2.26 (0.08)	0.58 (0.08)
2015	2.36 (0.08)	0.56 (0.08)

Tabela 4: Número de gols por partida e saldo médio por partida da Série A do Campeonato Brasileiro.

### Número médio de gols por partida - 1997 a 2015

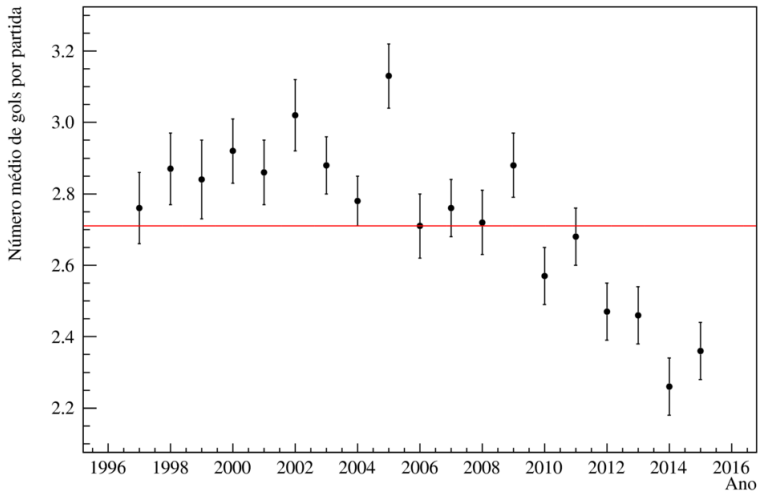


Figura 3: Número médio de gols por partida.  $\chi^2 = 126.1$ ,  $\nu = 18$

## Saldo de gols médio - 1997 a 2015

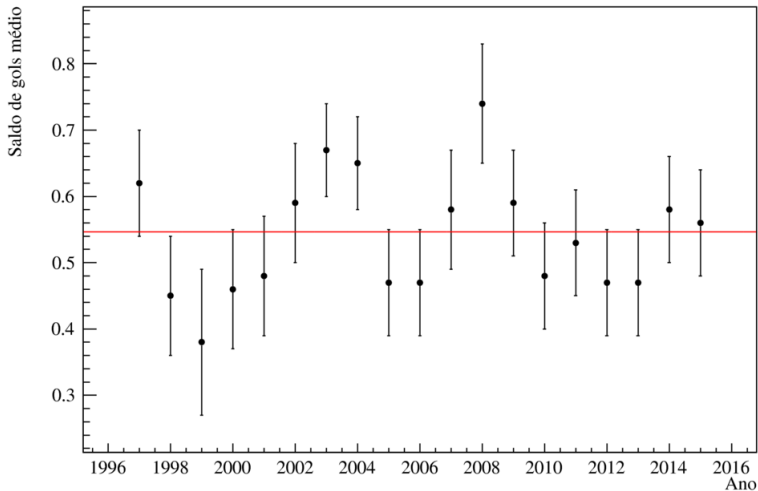


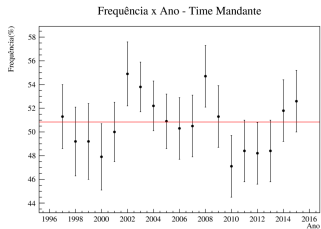
Figura 4: Saldo médio por partida.  $\chi^2 = 20.9, \nu = 18$

# Campeonato Brasileiro A

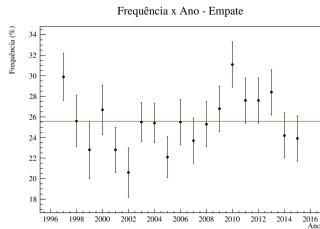
## Resultados

Brasileiro	Mandante	Empate	Visitante
1997	51.3 (2.7) %	29.9 (2.3) %	18.8 (2.3) %
1998	49.2 (2.9) %	25.6 (2.5) %	25.3 (2.5) %
1999	49.2 (3.2) %	22.8 (2.8) %	28.0 (2.7) %
2000	47.9 (2.8) %	26.7 (2.4) %	25.5 (2.3) %
2001	50.0 (2.5) %	22.8 (2.2) %	27.2 (2.2) %
2002	54.9 (2.7) %	20.6 (2.4) %	24.5 (2.3) %
2003	53.8 (2.1) %	25.5 (1.9) %	20.7 (1.8) %
2004	52.2 (2.1) %	25.4 (1.9) %	22.5 (1.8) %
2005	50.9 (2.3) %	22.1 (2.0) %	27.1 (2.0) %
2006	50.3 (2.6) %	25.5 (2.2) %	24.2 (2.2) %
2007	50.5 (2.6) %	23.7 (2.2) %	25.8 (2.2) %
2008	54.7 (2.6) %	25.3 (2.2) %	20.0 (2.2) %
2009	51.3 (2.6) %	26.8 (2.2) %	21.8 (2.2) %
2010	47.1 (2.6) %	31.1 (2.2) %	21.8 (2.2) %
2011	48.4 (2.6) %	27.6 (2.2) %	23.9 (2.2) %
2012	48.2 (2.6) %	27.6 (2.2) %	24.2 (2.2) %
2013	48.4 (2.6) %	28.4 (2.2) %	23.2 (2.2) %
2014	51.8 (2.6) %	24.2 (2.2) %	23.9 (2.2) %
2015	52.6 (2.6) %	23.9 (2.2) %	23.4 (2.2) %

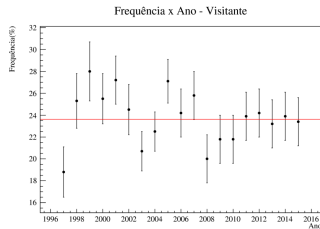
Tabela 5: Relação entre vitórias, empates e derrotas para a Série A do Campeonato Brasileiro.



(a)  $\chi^2 = 14.2, \nu = 18$



(b)  $\chi^2 = 25.4, \nu = 18$

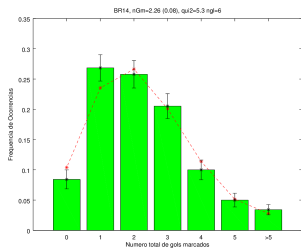


(c)  $\chi^2 = 22.3, \nu = 18$

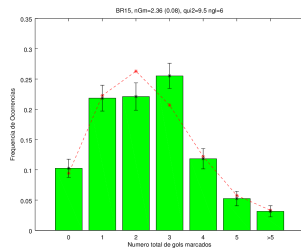
Figura 5: Relação entre vitórias, empates e derrotas para a Série A do Campeonato Brasileiro entre 1997 e 2015.

# Campeonato Brasileiro A

## Distribuição de gols por partida



(a) Campeonato BR 2014



(b) Campeonato BR 2015

Figura 6: Número de gols por partida

PARECEM SEGUIR UMA POISSON! :o



# Detalhes Teóricos

## Hipóteses

- Pensando na posse de bola como fator central, a probabilidade  $p$  da posse resultar em ataque e do ataque resultar em gol é pequena. Sob esse ponto de vista a distribuição do número de gols em uma partida pode ser pensada como uma **Poisson** (lembrando que ataques são independentes).
- Campeonato por **pontos corridos**.

# Detalhes Teóricos

## O modelo

Em uma partida entre um time  $i$  (anfitrião) e outro  $j$  (visitante), espera-se que o placar seja

$$(x_{ij}, y_{ij})$$

de modo que

- $x_{ij}$  uma variável aleatória que segue uma distribuição Poisson de média  $\alpha_i$   $\beta_j$
- $y_{ij}$  uma variável aleatória que segue uma distribuição Poisson de média  $\gamma_i$   $\delta_j$

# Detalhes Teóricos

## O modelo II

tais que:

- $x_{ij}$  gols marcados pelo  $i$
- $y_{ij}$  gols marcados pelo  $j$
- $\alpha_k$  qualidade de ataque de um time  $k$  jogando em casa
- $\beta_k$  fraqueza de defesa de um time  $k$  jogando em fora de casa
- $\gamma_k$  fraqueza de defesa de um time  $k$  jogando em casa
- $\delta_k$  qualidade de ataque de um time  $k$  jogando em fora de casa

Podemos então escrever a distribuição de Poisson para  $x_{ij}$ :

$$P_{\alpha_i \beta_j}(x_{ij}) = \frac{e^{-\alpha_i \beta_j} (\alpha_i \beta_j)^{x_{ij}}}{x_{ij}!}$$

e podemos escrever a função verossimilhança:

$$L(\alpha, \beta) = \prod_i \prod_{j \neq i} \frac{e^{-\alpha_i \beta_j} (\alpha_i \beta_j)^{x_{ij}}}{x_{ij}!}$$

# Detalhes Teóricos

## O modelo III

Então é possível estimar esses parâmetros, característicos de cada time, a partir dos gols do campeonato, usando o método da máxima verossimilhança.

Da verossimilhança logarítmica, dados dois times  $i$  e  $j$

$$\ln L(\alpha, \beta) = \sum_i \sum_{j \neq i} (-\alpha_i \beta_j + x_{ij} \ln(\alpha_i \beta_j) - \ln(x_{ij}!))$$

$$\frac{\partial \ln L}{\partial \alpha_i} = \sum_{j \neq i} \left( -\tilde{\beta}_j + \frac{x_{ij}}{\tilde{\alpha}_i} \right) = 0$$

# Detalhes Teóricos

## O modelo IV

Então, para obtenção do parâmetro  $\bar{\alpha}_i$ , por exemplo:

$$\bar{\alpha}_i = \frac{\sum_{j \neq i} x_{ij}}{\sum_{j \neq i} \bar{\beta}_j} \quad \left( \text{chute : } \tilde{\beta}_j = \frac{\sum_{i \neq j} x_{ij}}{\sqrt{\sum_i \sum_{j \neq i} x_{ij}}} \right)$$

De modo geral, para obter apenas um conjunto de parâmetros relacionado a uma distribuição de gols, as seguintes restrições devem ser impostas:

$$\sum_i \alpha_i = \sum_i \beta_i$$

$$\sum_i \gamma_i = \sum_i \delta_i$$

Podemos calcular a incerteza dos parâmetros usando a Propagação de Incertezas. Para o parâmetro  $\tilde{\alpha}_i$ , por exemplo:

$$\sigma_{\tilde{\alpha}_i} = \sqrt{\sum_i \left( \frac{\partial \tilde{\alpha}_i}{\partial x_{ij}} \sigma_{x_{ij}} \right)^2}$$

Manipulando algebricamente, chegamos que a expressão para o cálculo da incerteza para o parâmetro  $\tilde{\alpha}_i$  é:

$$\sigma_{\tilde{\alpha}_i} = \sqrt{\frac{\tilde{\alpha}_i}{\sum_j \tilde{\beta}_j}}$$

# Detalhes Teóricos

## Observação

Table 2. Hierarchy of models, with changes in the value of twice the maximised log likelihood shown for Division I 1971-1972

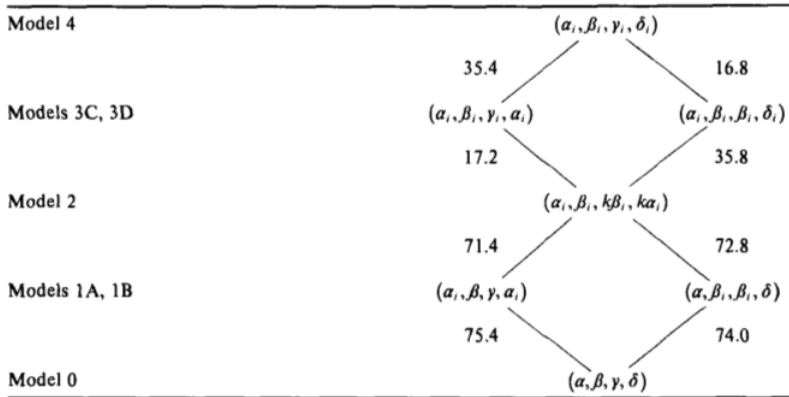


Figura 7: Níveis de hierarquia dos modelos de Maher.



# Implementação

- Usamos um conjunto de scripts escritos no Octave/Matlab. Esses scripts são capazes de obter os parâmetros  $\alpha$ ,  $\beta$ ,  $\gamma$  e  $\delta$  adotando os modelos 0, 2 e 4.
- Focamos a análise nas edições de 2014 e 2015 da Série A do Campeonato Brasileiro e utilizamos o modelo 4 do Método de Maher.

# Resultados

## Teste inicial

Table 1. Maximum likelihood estimates of the parameters for Division 1 1971–1972.

	home attack $\alpha$	away defence $\beta$	home defence $\gamma$	away attack $\delta$
Arsenal	1.36	1.03	0.64	1.06
Chelsea	1.55	1.18	0.97	0.83
Coventry City	1.05	1.66	1.12	0.84
Crystal Palace	0.99	1.28	1.49	0.65
Derby County	1.62	0.89	0.50	1.24
Everton	1.06	1.17	0.81	0.44
Huddersfield Town	0.46	1.37	1.06	0.74
Ipswich Town	0.72	1.27	0.93	0.98
Leeds United	2.02	0.82	0.49	0.91
Leicester City	0.69	1.31	0.54	1.10
Liverpool	1.78	0.54	0.78	0.78
Manchester City	1.82	1.17	0.75	1.40
Manchester United	1.49	1.35	1.31	1.49
Newcastle United	1.14	1.29	0.88	0.93
Nottingham Forest	0.98	1.96	1.43	1.10
Sheffield United	1.49	1.31	1.28	1.09
Southampton	1.21	1.98	1.38	1.05
Stoke City	0.99	1.17	1.20	0.64
Tottenham Hotspur	1.71	1.12	0.63	0.87
West Bromwich Albion	0.84	1.16	1.13	0.99
West Ham United	1.18	1.22	0.92	0.78
Wolverhampton Wanderers	1.34	1.30	1.15	1.48

Figura 8: Número médio de gols e saldo médio por partida.

# Resultados

## Teste inicial

Equipes	$\alpha$	$\beta$	$\gamma$	$\delta$
Arsenal	1,36(22)	1,03(19)	0,64(17)	1,06(22)
Chelsea	1,55(24)	1,12(20)	0,97(21)	0,83(20)
Conventry City	1,05(19)	1,66(25)	1,12(23)	0,84(20)
Crystal Palace	0,99(19)	1,28(21)	1,49(26)	0,65(17)
Derby County	1,62(24)	0,89(18)	0,50(15)	1,24(24)
Everton	1,06(20)	1,17(20)	0,81(19)	0,44(14)
Huddersfield Town	0,46(13)	1,37(22)	1,06(22)	0,74(19)
Ipswich Town	0,72(16)	1,27(21)	0,93(21)	0,98(21)
Leeds United	2,02(27)	0,82(17)	0,49(15)	0,91(21)
Leicester City	0,69(16)	1,30(22)	0,54(16)	1,10(23)
Liverpool	1,78(25)	0,54(14)	0,78(19)	0,78(19)
Manchester City	1,82(26)	1,17(21)	0,75(19)	1,40(26)
Manchester United	1,49(23)	1,35(22)	1,31(25)	1,49(26)
Newcastle United	1,14(20)	1,29(22)	0,88(20)	0,93(21)
Nottingham Forest	0,98(19)	1,96(27)	1,43(26)	1,10(23)
Sheffield United	1,49(23)	1,31(22)	1,28(24)	1,09(23)
Southampton	1,21(21)	1,98(27)	1,38(25)	1,05(22)
Stoke City	0,99(19)	1,17(21)	1,20(24)	0,64(17)
Tottenham Hotspur	1,71(25)	1,12(20)	0,63(17)	0,87(20)
West Bromwich Albion	0,84(17)	1,16(21)	1,13(23)	0,99(21)
West Ham United	1,18(21)	1,22(21)	0,92(21)	0,78(19)
Wolverhampton Wanderers	1,33(22)	1,30(22)	1,15(23)	1,48(26)

Tabela 6: Parâmetros gerados pelo script.

# Resultados

## Série A do Campeonato Brasileiro

Brasileiro 2014 - Times	$\alpha$	$\beta$	$\gamma$	$\delta$
Cruzeiro	1,88(28)	0,96(20)	1,00(23)	1,38(27)
São Paulo	1,41(24)	1,07(21)	0,95(23)	1,55(29)
Internacional	1,63(26)	1,13(22)	0,92(22)	0,92(22)
Corinthians	1,38(24)	0,71(17)	0,86(22)	0,97(23)
Atlético Mineiro	1,23(23)	1,15(22)	0,70(20)	1,30(27)
Fluminense	1,76(27)	1,04(21)	1,11(25)	1,22(26)
Grêmio	1,03(21)	0,57(15)	0,62(18)	0,68(19)
Atlético Paranaense	1,06(21)	1,23(23)	0,81(21)	1,08(24)
Santos	1,10(21)	1,01(21)	0,69(19)	0,96(23)
Flamengo	1,29(23)	1,42(24)	0,86(22)	0,97(23)
Sport	1,02(21)	1,31(23)	0,91(22)	0,75(20)
Goiás	1,32(24)	1,16(22)	0,78(21)	0,46(16)
Figueirense	1,06(21)	1,19(22)	1,14(25)	0,76(20)
Coritiba	1,07(21)	1,49(25)	0,63(19)	1,02(24)
Chapecoense	1,07(21)	1,45(25)	0,63(19)	0,85(21)
Palmeiras	0,94(20)	1,57(26)	1,31(27)	0,76(20)
Vitória	1,03(21)	1,40(24)	1,25(26)	0,82(21)
Bahia	0,62(16)	1,12(22)	0,98(23)	0,98(23)
Botafogo	1,06(21)	1,27(23)	1,06(24)	0,40(15)
Criciúma	0,85(19)	1,57(26)	1,12(25)	0,52(17)

Tabela 7: Parâmetros para o Campeonato Brasileiro de 2014.

# Resultados

## Série A do Campeonato Brasileiro

Brasileiro 2015 - Times	$\alpha$	$\beta$	$\gamma$	$\delta$
Corinthians	1,76(27)	0,89(19)	0,63(18)	1,64(29)
Atlético Mineiro	1,57(25)	1,20(22)	1,15(25)	1,62(29)
Grêmio	1,50(25)	0,75(18)	0,83(21)	0,94(22)
São Paulo	1,54(25)	1,37(24)	0,89(22)	1,00(23)
Internacional	1,22(22)	1,26(23)	0,49(16)	0,59(18)
Sport	1,44(24)	1,28(23)	0,50(16)	1,08(24)
Santos	2,04(29)	1,18(22)	0,82(21)	0,66(19)
Cruzeiro	1,21(22)	1,05(21)	0,61(18)	0,87(21)
Palmeiras	1,35(24)	1,23(23)	1,33(26)	1,64(29)
Atlético Paranaense	1,31(23)	1,23(23)	1,10(24)	0,73(20)
Ponte Preta	0,95(20)	1,08(21)	0,84(21)	1,05(23)
Flamengo	1,22(22)	1,18(22)	1,44(28)	0,97(23)
Fluminense	1,09(21)	1,30(23)	1,05(23)	0,84(21)
Chapecoense	1,00(20)	1,08(21)	1,03(23)	0,61(18)
Coritiba	0,65(16)	1,19(22)	0,77(20)	0,88(22)
Figueirense	0,79(18)	1,41(24)	0,95(22)	1,00(23)
Avai	1,15(22)	1,48(25)	1,42(27)	0,68(19)
Vasco	0,57(15)	1,31(23)	1,27(26)	0,85(21)
Goiás	0,97(20)	1,38(24)	0,94(22)	0,94(22)
Joinville	0,83(19)	1,33(23)	0,91(22)	0,39(14)

Tabela 8: Parâmetros para o Campeonato Brasileiro de 2015.

# Resultados

## Série A do Campeonato Brasileiro

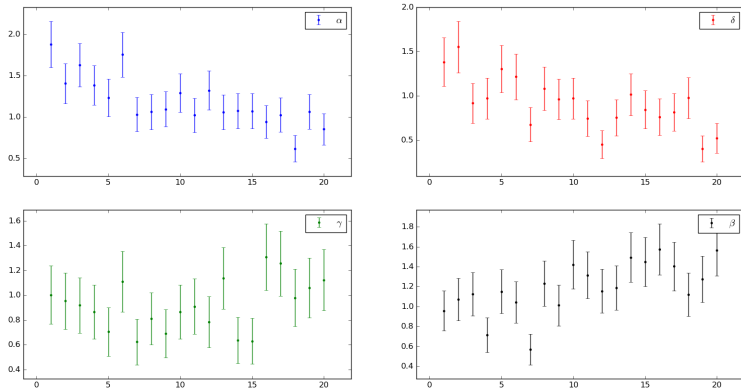


Figura 9: Parâmetros gerados para os times do Campeonato Brasileiro da Série A de 2014.

# Resultados

## Série A do Campeonato Brasileiro

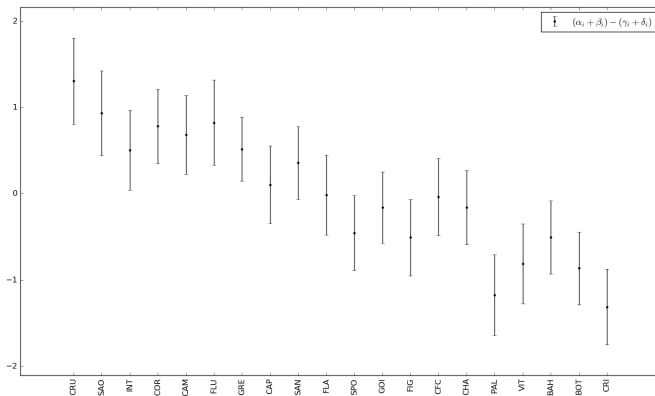
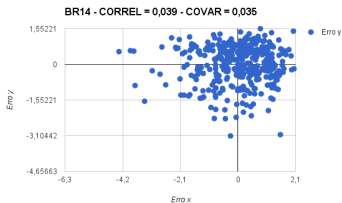
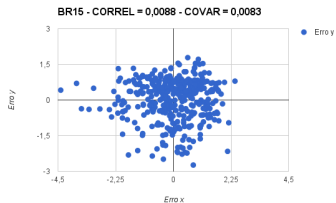


Figura 10: Parâmetro Soma para qualificar os times do Campeonato Brasileiro da Série A de 2014.

# Resultados



(a) Campeonato BR 2014



(b) Campeonato BR 2015

Figura 11: Dispersão de Erros - Esperado x Observado



# Resultados

## Série A do Campeonato Brasileiro

	Modelo 0	Modelo 2	Modelo 4
2014	46*	56	63
2015	33*	43	54

Tabela 9: Acertos.

# Resultados

## Série A do Campeonato Brasileiro

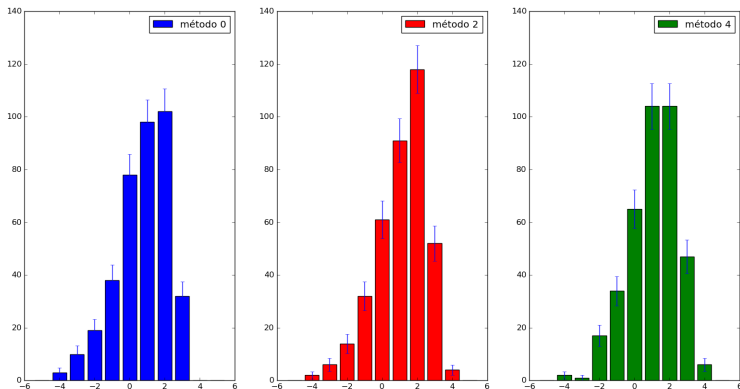


Figura 12: Excesso de contagem de gols do Modelo de Maher para os níveis de hierarquia 0, 2 e 4

# Conclusões

- As variáveis  $\alpha$ ,  $\beta$ ,  $\gamma$  e  $\delta$  tem incertezas grandes demais para poderem ser usadas como parâmetros de classificação qualitativa.
- A hipótese de que o número total de gols marcados numa partida, por exemplo, seja uma Poisson foi verificada.
- Os diferentes níveis de hierarquia influenciam o resultado final.

VALEU ZWINGLIO!

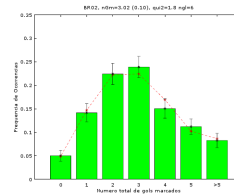
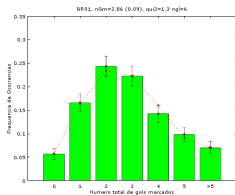
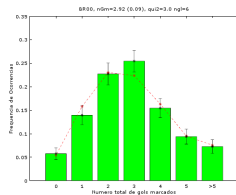
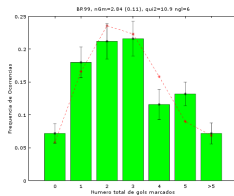
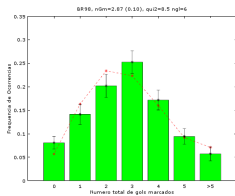
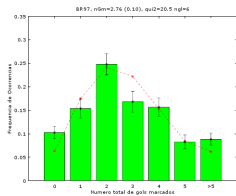


Figura 13: Número médio de gols de 1997 a 2002.

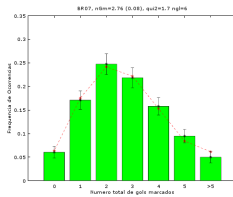
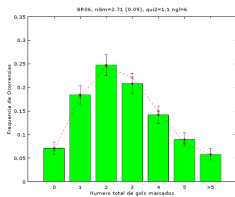
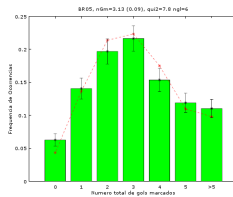
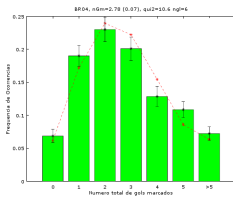
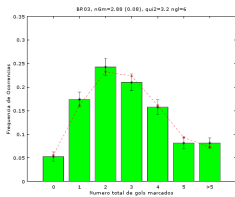


Figura 14: Número médio de gols de 2003 a 2007.

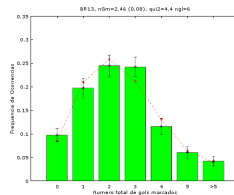
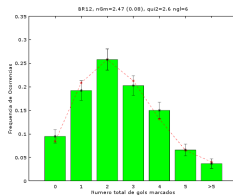
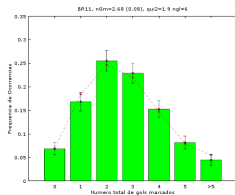
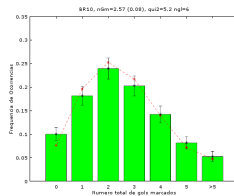
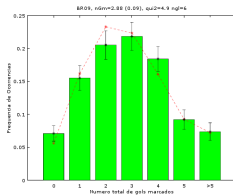
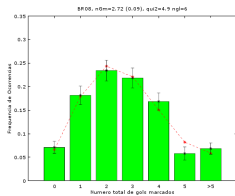


Figura 15: Número médio de gols de 2008 a 2013.

## Script 1 I

## script\_analise\_numero\_total\_de\_gols\_marcados\_corrigido.m

```

1 %%
2 clear
3 Ncamp= 19;
4 lista=zeros(1,Ncamp);
5 NOME_XLS = 'BRASILEIRO.xls';
6 for JOTA = Ncamp-1
7 [Nome, Placar,ok,okM, Descricao ] = carrega_tabela( NOME_XLS, JOTA );
8 ntG = Placar(:,1) + Placar(:,2);
9 ntGm = mean( ntG );
10 lista(1,JOTA)=ntGm;
11
12 htG = hist( ntG, 0:20 );
13 N = sum( htG );
14 ftG = htG / N;
15 ftG_pois = poisspdf(0:20,ntGm);
16
17 nMAX = 7;
18 ftG(nMAX) = sum( ftG(nMAX:end) ); ftG((nMAX+1):end) = [];
19 ftG_pois(nMAX) = sum( ftG_pois(nMAX:end) ); ftG_pois((nMAX+1):end) = [];
20
21 sftG = sqrt( ftG_pois.*(1-ftG_pois)/N );
22 Qui2 = sum( ((ftG-ftG_pois)./sftG).^2 );
23
24 %
25 figure
26 bar( 0:(nMAX-1), ftG, 'g' )
27 hold on

```



## Script 1 II

script\_analise\_numero\_total\_de\_gols\_marcados\_corrigido.m

```

28 plot( 0:(nMAX-1), ftG_poisson, '*--r' )
29 errorbar( 0:(nMAX-1), ftG, sftG, '*k' )
30 xlim( [-1 nMAX] )
31 xlabel( 'Numero total de gols marcados' )
32 ylabel( 'Frequencia de Ocorrencias' )
33
34 title( sprintf( '%s, nGm=%.2f (%.2f), qui2=%.1f ngl=%d', ...
35             Descricao, ntGm, std(ntG)/sqrt(length(ntG)), Qui2, nMAX-1 ) )
36
37 fprintf( '%s, nGm=%.2f (%.2f), qui2=%.4f ngl=%d\n', ...
38         Descricao, ntGm, std(ntG)/sqrt(length(ntG)), Qui2, nMAX-1 )
39
40 for n=0:(nMAX-1)
41     XLB{n+1}=sprintf( '%d',n);
42 end
43 XLB{end} = ['>' XLB{end-1}];
44 set( gca, 'xtick', (0:(nMAX-1)), 'xticklabel', XLB )
45 end
46 fprintf('ntgmM = %.2f(%.2f)', mean(lista), std(lista))

```

## Script 2 I

## script\_saldo\_de\_gols\_corrigido.m

```

1 %%
2 clear
3 Ncamp= 19;
4 lista=zeros(1,Ncamp);
5 for JOTA = Ncamp-1
6 [Nome, Placar ,ok,okM, Descricao ] = carrega_tabela( 'BRASILEIRO.xls', JOTA );
7 SALDO = Placar(:,1) - Placar(:,2);
8 SALDOm = mean( SALDO );
9 sSALDO = std( SALDO );
10 lista(1,JOTA)= SALDOm;
11
12 bins = (-3:4);
13 hSALDO = hist( SALDO, bins );
14 binsINTERP = (bins(2)-.4):.1:(bins(end)-1)+.4);
15 hSALDO_GAUSS = length(SALDO)*normpdf(binsINTERP ,SALDOm,sSALDO);
16 %
17 ok_borda = zeros(size(hSALDO)); ok_borda([1 end])=1;
18 figure
19 bar( bins , hSALDO.*ok_borda , 'c' , .5 )
20 hold on
21 bar( bins , hSALDO.*(1-ok_borda) , 'y' )
22 plot( binsINTERP , hSALDO_GAUSS , '*-r' )
23 errorbar( bins , hSALDO , sqrt( hSALDO.*(1-hSALDO)/length(SALDO))), '*k' )
24 xlim( [bins(1)-.5 bins(end)+.5] )
25 set( gca , 'fontsize', 12, 'tickdir', 'out' )
26 xlabel( 'Saldo de gols (marcados-sofridos)', 'fontsize', 14 )
27 ylabel( 'Numero de Ocorrencias' )

```

# Script 2 II

script\_saldo\_de\_gols\_corrigeo.m

```
28 title( sprintf( '%s, Sm=%.2f (%.2f)', Descricao, SALDOm, sSALDO/sqrt(length(
    SALDO)) ) )
29 fprintf( '%s, Sm=%.2f (%.2f)\n', Descricao, SALDOm, sSALDO/sqrt(length(SALDO)) )
30 end
31
32 fprintf('Saldo_medio_camp = %.2f(%.2f)', mean(lista), std(lista))
```

# Script 3 I

## script\_f2016\_ajuste\_Maher\_Novo.m

```
1 function [alfa , G, A, gamma, PlacarFIT , PARAM] = f2016_ajuste_Maher( Placar ,
    indices , Modelo, PARAMini )
2 % [alfa , G, A, gamma, PlacarFIT] = f2016_ajuste_Maher( Placar , indices , Modelo ,
    PARAMini )
3 % ou
4 % [alfa , G, A, gamma, PlacarFIT] = f2016_ajuste_Maher( Placar , Times, Modelo ,
    PARAMini )
5
6 if ~exist('Modelo','var') || isempty(Modelo)
7     Modelo = 2; % O modelo padrao eh o 2
8 end
9
10 if iscell(indices)
11     Times = indices;
12     [Equipes , indices] = identifique_times( Times );
13 end
14
15 x = Placar(:,1);
16 y = Placar(:,2);
17 ind1 = indices(:,1);
18 ind2 = indices(:,2);
19 Ntimes = max(indices(:));
20 Njogos = length(x);
21
22 Sx = sum( x );
23 Sy = sum( y );
24 alfa = zeros( Ntimes , 1 );
```

## Script 3 II

script\_f2016\_ajuste\_Maher\_Novo.m

```

25 G = zeros( Ntimes, 1 );
26 A = zeros( Ntimes, 1 );
27 gamma = zeros( Ntimes, 1 );
28 iter = 1;
29 MaxIter = 50;
30 CONVERGIU = 0;
31 switch( Modelo )
32     case{4}
33         if ~exist( 'PARAMini','var' ) || isempty(PARAMini)
34             for i=1:Ntimes
35                 alfa(i) = sum( x(ind1==i) )/sqrt(Sx);
36                 gamma(i) = sum( y(ind1==i) )/sqrt(Sy);
37                 A(i) = sum( y(ind2==i) )/sqrt(Sy);
38                 G(i) = sum( x(ind2==i) )/sqrt(Sx);
39             end
40         else
41             if size(PARAMini,1)==4*Ntimes
42                 PARAMini = PARAMini(:,end);
43             end
44             alfa = PARAMini( 0*Ntimes + (1:Ntimes) );
45             G = PARAMini( 1*Ntimes + (1:Ntimes) );
46             A = PARAMini( 2*Ntimes + (1:Ntimes) );
47             gamma = PARAMini( 3*Ntimes + (1:Ntimes) );
48         end
49         [ alfa , G, A, gamma ] = normaliza_soma( alfa , G, A, gamma );
50         PARAM(:,1) = [alfa; G; A; gamma];
51         while( (CONVERGIU==0) && (iter<MaxIter) )
52             for i=1:Ntimes

```

## Script 3 III

script\_f2016\_ajuste\_Maher\_Novo.m

```

53         alfa(i) = sum( x(ind1==i) )/sum(G(ind2(ind1==i)));
54         gamma(i) = sum( y(ind1==i) )/sum(A(ind2(ind1==i)));
55     end
56     [ alfa , G, A, gamma ] = normaliza_soma( alfa , G, A, gamma );
57
58     for j=1:Ntimes
59         G(j) = sum( x(ind2==j) )/sum( alfa(ind1(ind2==j)));
60         A(j) = sum( y(ind2==j) )/sum(gamma(ind1(ind2==j)));
61     end
62     [ alfa , G, A, gamma ] = normaliza_soma( alfa , G, A, gamma );
63
64     PARAM(:, iter+1)=[alfa; G; A; gamma];
65     if sum(abs(PARAM(:, iter)-PARAM(:, iter+1)))<1e-3
66         CONVERGIU = 1;
67     else
68         iter = iter + 1;
69     end
70 end
71 case{2}
72     k2 = Sy/Sx;
73     if ~exist( 'PARAMini', 'var' ) || isempty(PARAMini)
74         for i=1:Ntimes
75             alfa(i) = (sum(x(ind1==i))+sum(y(ind2==i))) / sqrt(Sx+Sy);
76             G(i) = (sum(x(ind2==i))+sum(y(ind1==i))) / sqrt(Sx+Sy);
77         end
78     else
79         if size(PARAMini,1)==4*Ntimes || size(PARAMini,1)==2*Ntimes
80             PARAMini = PARAMini(:, end);

```

## Script 3 IV

script\_f2016\_ajuste\_Maher\_Novo.m

```

81         end
82         alfa = PARAMini( 0*Ntimes + (1:Ntimes) );
83         G = PARAMini( 1*Ntimes + (1:Ntimes) );
84     end
85     [ alfa , G ] = normaliza_soma( alfa , G );
86     PARAM(:,1) = [ alfa; G; sqrt(k2)*alfa; sqrt(k2)*G];
87     while( (CONVERGIU==0) && (iter<MaxIter) )
88         for i=1:Ntimes
89             alfa(i) = (sum(x(ind1==i))+sum(y(ind2==i))) / ...
90                 ( sum(G(ind2(ind1==i))) + k2*sum(G(ind1(ind2==i)))) );
91         end
92         [ alfa , G ] = normaliza_soma( alfa , G );
93
94         for j=1:Ntimes
95             G(j) = (sum(y(ind1==j))+sum(x(ind2==j))) / ...
96                 ( k2*sum(alfa(ind1(ind2==j))) + sum(alfa(ind2(ind1==j)))) );
97         end
98         [ alfa , G ] = normaliza_soma( alfa , G );
99         PARAM(:, iter+1) = [ alfa; G; sqrt(k2)*alfa; sqrt(k2)*G];
100         if sum(abs(PARAM(:, iter)-PARAM(:, iter+1)))<1e-3
101             CONVERGIU = 1;
102         else
103             iter = iter + 1;
104         end
105     end
106     A = sqrt(k2)*alfa;
107     gamma = sqrt(k2)*G;
108 case {0}

```

## Script 3 V

## script\_f2016\_ajuste\_Maher\_Novo.m

```

109     alfa (:) = sqrt( Sx/Njogos );
110     G(:) = sqrt( Sx/Njogos );
111     A(:) = sqrt( Sy/Njogos );
112     gamma (:) = sqrt( Sy/Njogos );
113     PARAM = [ alfa ; G ; A ; gamma ];
114 end
115 PlacarFIT = [ alfa(ind1).*G(ind2)    A(ind2).*gamma(ind1) ];
116
117 function [ alfa , G, A, gamma ] = normaliza_soma( alfa , G, A, gamma )
118     SOMA_alfa = sum( alfa );
119     SOMA_G = sum( G );
120     alfa = alfa*sqrt( SOMA_G / SOMA_alfa );
121     G = G*sqrt( SOMA_alfa / SOMA_G );
122     if nargin==4
123         SOMA_A = sum( A );
124         SOMA_gamma = sum( gamma );
125         A = A*sqrt( SOMA_gamma / SOMA_A );
126         gamma = gamma*sqrt( SOMA_A / SOMA_gamma );
127     end
128 end
129 return
130 %% Exemplo de uso
131 clear
132 [Nome , Placar, ok, okM, QUAL ] = carrega_tabela('Brasileiro.xls',13);
133 [Eq, ind] = identifique_times( Nome );
134 [alfa , G, A, gamma, PlacarFIT , PARAM] = f2016_ajuste_Maher( Placar , ind , 4 );
135 [alfa_O , G_O, A_O, gamma_O, PlacarFIT_O , PARAM_O] = f2016_ajuste_Maher_OLD(
    Placar , ind , 4 );

```



# Script 3 VI

script\_f2016\_ajuste\_Maher\_Novo.m

```
136 %[alfa4 , G4, A4, gamma4, PlacarFIT4 , PARAM4] = f2016_ajuste-Maher( Placar , ind ,  
137     4 );  
137 end
```

## Script 4 I

## script\_testa\_maher.m

```

1  clear
2  for JOTA = 1:19
3  [Nome , Placar , ok , okM , QUAL ] = carrega_tabela('BRASILEIRO.xls',JOTA);
4  [Eq, ind] = identifique_times( Nome );
5  [alfa , G, A, gamma, PlacarFIT] = f2016_ajuste_Maher( Placar , ind , 2 );
6  [alfa4 , G4, A4, gamma4, PlacarFIT4] = f2016_ajuste_Maher( Placar , ind , 4 );
7  [alfa0 , G0, A0, gamma0, PlacarFIT0] = f2016_ajuste_Maher( Placar , ind , 0 );
8
9  Y = Placar(:);
10 F = PlacarFIT(:);
11 si = sqrt(PlacarFIT4(:));
12 qui2 = sum( ((Y-F)./si).^2 );
13 %fprintf( '%s %2d q2=%6.1f ngl=%d\n', QUAL, JOTA, qui2, length(Y)-2*max(ind(:)) )
14 fprintf( '%s %2d q2=%6.1f ngl=%d\n', QUAL, JOTA, qui2, length(Y)-4*max(ind(:))+2
15 )
15 end
16
17 %%
18 PlacarFIT_original = PlacarFIT;
19 for qMC = 1e2:-1:1
20 PlacarMC = poissrnd(PlacarFIT_original);
21 [alfaMC(:,qMC), GMC(:,qMC), AMC(:,qMC), gammaMC(:,qMC), PlacarFIT] =
    f2016_ajuste_Maher( PlacarMC , ind , 4 );
22 end

```