

UNIVERSIDADE FEDERAL FLUMINENSE  
CENTRO TECNOLÓGICO  
MESTRADO EM ENGENHARIA DE PRODUÇÃO

ALESSANDRO MARTINS ALVES

Previsão de Resultados para o Campeonato Brasileiro de Futebol com Base no Modelo Logito  
Ordinal

Niterói  
2009

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

ALESSANDRO MARTINS ALVES

Previsão de Resultados para o Campeonato Brasileiro de Futebol com Base no Modelo Logito  
Ordinal

Dissertação apresentada ao Curso de Mestrado em Engenharia de Produção da Universidade Federal Fluminense, como requisito parcial para obtenção do Grau de Mestre. Área de concentração: Sistemas, Apoio à Decisão e Logística.

Orientador: Prof. Dr. JOÃO CARLOS CORREA BAPTISTA SOARES DE MELLO

Niterói  
2009

ALESSANDRO MARTINS ALVES

PREVISÃO DE RESULTADOS PARA O CAMPEONATO BRASILEIRO DE FUTEBOL  
COM BASE NO MODELO LOGITO ORDINAL

Dissertação apresentada ao Curso de Mestrado em Engenharia de Produção da Universidade Federal Fluminense, como requisito parcial para obtenção do Grau de Mestre. Área de concentração: Sistemas, Apoio à Decisão e Logística.

Aprovado em 25 de setembro de 2009

BANCA EXAMINADORA

---

Prof. Dr. João Carlos Correia Baptista Soares de Mello - Orientador  
Universidade Federal Fluminense

---

Prof. Dr. Annibal Parracho Sant'Anna  
Universidade Federal Fluminense

---

Dr. Luiz Biondi Neto  
Universidade do Estado do Rio de Janeiro

Niterói  
2009

## **AGRADECIMENTOS**

A Deus, pela força em todos os momentos difíceis!

Ao meu orientador, Prof. João Carlos, pelos ensinamentos de teoria e pela confiança em orientar uma dissertação fora de sua área de atuação.

A minha esposa, por entender minha ausência.

A Thiago Ramos pela ajuda futebolística e pelo companheirismo.

A todos os outros colegas de mestrado, que de alguma contribuíram para o meu aprendizado através do compartilhamento de experiências.

A todos os amigos que acabaram por entender o meu distanciamento e apoiaram-me em todos os momentos.

*“Quem quiser vencer  
na vida deve fazer como  
seus sábios:  
mesmo com a alma partida,  
ter um sorriso nos lábios ”*

Dinamor

## SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>13</b>
1.1 Objetivo .....	16
1.2 Estrutura do Trabalho .....	17
<b>2. CAMPEONATO BRASILEIRO DE FUTEBOL .....</b>	<b>18</b>
2.1 Evolução do Campeonato Brasileiro.....	18
2.2 Formato de Disputa .....	26
2.3 Mudanças de Regra e Formato dos Campeonatos Brasileiros .....	28
<b>3. REVISÃO BIBLIOGRÁFICA .....</b>	<b>33</b>
<b>4. METODOLOGIA.....</b>	<b>36</b>
4.1- Modelos Lineares Generalizados (MLG) .....	36
4.1.1- Modelo de Regressão Logística Binária .....	38
4.1.2- Modelo Logístico Simples .....	40
4.1.3- Modelo Logístico Múltiplo.....	42
4.1.4- Testes e Validação do Modelo .....	45
4.2 – Simulações .....	48
<b>5. MODELOS DO PRESENTE TRABALHO.....</b>	<b>52</b>
5.1 – Descrição dos Modelos .....	53
5.1.1 – Modelo 1 – Força do Time e Vantagem de Jogar em Casa .....	54
5.1.2 – Modelo 2 - Força do Time, Vantagem de Jogar em Casa e Desempenho Recente Dentro e Fora de Casa .....	56
5.1.3 – Modelo 3 - Força do Time, Vantagem de Jogar em Casa e Desempenho Recente do Time.....	57
<b>6. ANÁLISE DE RESULTADOS.....</b>	<b>59</b>
6.1 – Análise Exploratória dos jogos.....	59
6.2 – Análise dos modelos.....	62
6.2.1 – Modelo 1 - Força do Time, Vantagem de Jogar em Casa .....	63

<b>6.2.2 – Modelo 2 - Força do Time, Vantagem de Jogar em Casa e Desempenho Recente Dentro e Fora de Casa .....</b>	<b>71</b>
<b>6.2.3 – Modelo 3 - Força do Time, Vantagem de Jogar em Casa e Desempenho Recente do Time.....</b>	<b>75</b>
<b>6.2.4 – Comparação dos modelos .....</b>	<b>79</b>
<b>7. CONCLUSÕES.....</b>	<b>81</b>
<b>8. REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>83</b>
<b>9. ANEXOS .....</b>	<b>87</b>



## LISTA DE TABELAS

<b>TABELA 2.1 – CAMPEÕES DA TAÇA BRASIL .....</b>	<b>18</b>
<b>TABELA 2.2 – CAMPEÕES DA TAÇA DE PRATA.....</b>	<b>19</b>
<b>TABELA 2.3 – TIMES QUE PARTICIPARAM DA COPA UNIÃO DE 1987 .....</b>	<b>20</b>
<b>TABELA 2.4 – NÚMERO DE EQUIPES QUE CONQUISTARAM O TÍTULO EM CAMPEONATOS PELO MUNDO .....</b>	<b>22</b>
<b>TABELA 2.4 – MAIORES GOLEADAS DO CAMPEONATO BRASILEIRO ATÉ 2008 .....</b>	<b>26</b>
<b>TABELA 2.5 – NÚMERO DE EQUIPES DISPUTANDO O CAMPEONATO ATÉ 2008 .....</b>	<b>28</b>
<b>TABELA 4.1: CRUZAMENTO ENTRE A DISTRIBUIÇÃO DA VARIÁVEL RESPOSTA E DA FUNÇÃO DE LIGAÇÃO.....</b>	<b>38</b>
<b>TABELA 6.1: PONTUAÇÃO TOTAL E % DOS PONTOS OBTIDOS COMO MANDANTE E COMO VISITANTE .....</b>	<b>61</b>
<b>TABELA 6.2.1: COEFICIENTES DO MODELO 1 PARA A RODADA 20 .....</b>	<b>63</b>
<b>TABELA 6.2.1.2: RESULTADOS PARA O TIME COM MANDO DE CAMPO ATÉ A RODADA 20.....</b>	<b>64</b>
<b>TABELA 6.6: NÚMERO DE TÉCNICOS QUE CADA EQUIPE TEVE NO CAMPEONATO DE 2008 .....</b>	<b>69</b>
<b>TABELA 6.7: TESTES – MODELO 1 .....</b>	<b>70</b>
<b>TABELA 6.11: TESTES – MODELO 2 .....</b>	<b>74</b>
<b>TABELA 6.15: TESTES – MODELO 3 .....</b>	<b>78</b>

## LISTA DE FIGURAS

<b>FIGURA 2.1 – NÚMERO DE VEZES QUE CADA EQUIPE FOI CAMPEÃ DO CAMPEONATO BRASILEIRO.....</b>	<b>21</b>
<b>FIGURA 2.2 – NÚMERO DE GOLS DOS MAIORES ARTILHEIROS DA HISTÓRIA DO CAMPEONATO BRASILEIRO.....</b>	<b>25</b>
<b>FIGURA 4.1 – FUNÇÃO SIGMÓIDE .....</b>	<b>39</b>
<b>FIGURA 6.1 – DISTRIBUIÇÃO DOS RESULTADOS DAS PARTIDAS PELO TIME QUE ESTÁ JOGANDO EM CASA.....</b>	<b>59</b>
<b>FIGURA 6.2 – DESEMPENHO DOS MANDANTES E VISITANTES .....</b>	<b>60</b>
<b>FIGURA 6.3 – FORÇA DOS TIMES - MODELO 1.....</b>	<b>65</b>
<b>FIGURA 6.4 – CHANCES DE CLASSIFICAÇÃO PARA A COPA LIBERTADORES DA AMÉRICA RODADA A RODADA - MODELO 1 .....</b>	<b>66</b>
<b>FIGURA 6.5 – CHANCES DE REBAIXAMENTO PARA A SEGUNDA DIVISÃO - MODELO 1.....</b>	<b>68</b>
<b>FIGURA 6.8 – FORÇA DOS TIMES - MODELO 2.....</b>	<b>71</b>
<b>FIGURA 6.9 – CHANCES DE CLASSIFICAÇÃO PARA A COPA LIBERTADORES DA AMÉRICA - MODELO 2 .....</b>	<b>72</b>
<b>FIGURA 6.10 – CHANCES DE REBAIXAMENTO PARA A SEGUNDA DIVISÃO - MODELO 2.....</b>	<b>73</b>
<b>FIGURA 6.12 – FORÇA DOS TIMES - MODELO 3 .....</b>	<b>75</b>
<b>FIGURA 6.13 – CHANCES DE CLASSIFICAÇÃO PARA A COPA LIBERTADORES DA AMÉRICA - MODELO 3 .....</b>	<b>76</b>
<b>FIGURA 6.14 – CHANCES DE REBAIXAMENTO PARA A SEGUNDA DIVISÃO - MODELO 3.....</b>	<b>77</b>
<b>FIGURA 6.16 – FORÇA DO SÃO PAULO NOS TRÊS MODELOS.....</b>	<b>79</b>

## RESUMO

O presente trabalho utiliza o modelo logito ordinal para calcular as probabilidades de cada um dos resultados possíveis (vitória, empate e derrota) para cada jogo não realizado no campeonato Brasileiro de 2008, com base nos jogos realizados. Foi estimado um modelo a cada rodada, considerando os dados até a rodada e prevendo as rodadas seguintes. Com base nas probabilidades derivadas de cada modelo, foram simulados 1.000 resultados possíveis e calculado o percentual de vezes que cada time aparecia entre os quatro primeiros colocados (classificados para a copa Libertadores da América do ano seguinte) e entre os quatro últimos colocados (rebaixados para a série B). Os modelos testados no presente trabalho consideram como variáveis preditivas a vantagem de se jogar em casa (*Home Advantage*), a força individual de cada time e o desempenho recente do time no campeonato.

**Palavras-Chave:** Home Advantage, Logit ordinal, Previsões em futebol.

## ABSTRACT

The present work uses the logit model to figure out the probabilities of each possible result (win, draw or lose) for each game not realized in the Brazilian Soccer Championship 2008, using the realized ones previously. It was estimated one model for each turn, taking into consideration the data until the current one and forecasting the following ones. Based on the results of each model, it was simulated 1.000 possible results and calculated the percentage of times that each team was displayed between the first four places in the ranking (classified for the Libertadores da América Championship) and the last four places (which were relegated to a lower division). The used models considered as independent variables the home advantage, the individual team strength and the recent performance of the team.

**Keywords:** Home Advantage, Ordinal Logit, Football Forecasts;

## **1. INTRODUÇÃO**

Futebol é um esporte popular e famoso não somente no Brasil, mas em todo o mundo. As vantagens do futebol sobre os demais esportes são: regras fáceis, baixo custo e ausência de necessidade de utilização de material caro (Íbsan, 2006).

O planejamento e avaliação do desempenho dos times em um campeonato são fundamentais para que se possa aprender com as deficiências e melhorá-las, tornando o time mais competitivo nas temporadas seguintes.

Segundo Roche (2002) “as organizações – assim como os indivíduos – se vêm obrigados, cada vez mais, a refletir sobre o caminho a tomar no futuro, sobre o que fazer ou qual direção tomar. Em outras palavras, vêm-se obrigadas a planejar”.

Associações esportivas que pretendem elaborar um plano estratégico necessitam compreender a relação entre os fatores que influenciem no resultado de um campeonato, ou seja, compreender características que podem influenciar na classificação final do time, para que, aprimorando as mesmas, consigam uma melhor classificação no campeonato subsequente.

Para compreender o desempenho de um time no campeonato Brasileiro de futebol, faz-se necessário a utilização de um histórico de desempenho do time nos últimos campeonatos. Gelade (2007) identificou que o talento disponível em cada time impacta diretamente na força do time, ou seja, a tradição do time melhora seu desempenho. Vale ressaltar que este estudo tinha como objeto as seleções nacionais, porém no presente trabalho serão avaliados clubes de futebol, onde a tradição é mais relevante ao observar os títulos

conquistados por estas equipes.

A previsão dos resultados no futebol possui um papel fundamental no planejamento e na tomada de decisões sobre a postura e evolução do time ao longo do campeonato. Nos últimos anos, muito se dedicou a previsão de resultados esportivos para o desenvolvimento de estratégias de apostas, como, por exemplo, Craig e Hall (1994); Lee (1997) estudaram o futebol inglês e Stefani e Clarke (1992), que estudaram o futebol australiano.

Scarf e Shi (2008) avaliaram a importância dos jogos na liga inglesa de 2001 a 2004; com base nas previsões, os times poderiam determinar o esforço que deveria ser investido nos jogos, pois, em campeonatos do tipo *round robin*, cada jogo é importante para o resultado final do campeonato e, dependendo da importância desse jogo no resultado final do campeonato para o time, talvez fosse melhor poupar determinados jogadores chave para outros campeonatos que o time estivesse disputando. A previsão da importância de cada jogo é fundamental também para as televisões escolherem quais jogos transmitir.

A previsão dos resultados consiste na tentativa de extrapolar o comportamento futuro a partir dos dados históricos já observados. Para efeito de previsão, é de suma importância identificar que variáveis ou informações serão relevantes e fazem-se necessárias nos modelos preditivos.

Nos últimos anos, diversos trabalhos foram desenvolvidos para verificar o *home advantage*, ou vantagem de jogar em casa em diversos esportes. Essa vantagem impacta no resultado final dos jogos, sendo confirmada por diversos autores, sendo aplicada tanto no futebol como no beisebol, voleibol, futebol americano, entre outros esportes (Courneya e Carron, 1992; Pollard, 2002). Além do efeito em partidas individuais, existe o mesmo efeito em prol do país sede de grandes eventos como nas olimpíadas de inverno e verão, onde existe uma melhora no rendimento do país se comparada as últimas participações olímpicas como em Balmer et al. (2001 e 2003).

Em estudos anteriores, como em Pollard (1986), observou-se que há um grande predomínio das equipes mandantes, o que pode ser comprovado através de um aproveitamento de cerca de 60% a 67% de pontos obtidos em casa.

Acrescenta-se ainda o aspecto de familiaridade com este campo e estádio, sendo afirmado por Courneya e Carron (1992), Pollard (2002) que há uma maior consciência do atleta da "casa", o que lhe permite uma orientação mais eficaz nas ações decorrentes e exigidas na partida.

Segundo Pollard (2006), a grande familiaridade de um time com o seu estádio faz

com que o mesmo leve vantagem, quando enfrenta um adversário em seus domínios, logo, espera-se um desempenho melhor em casa do que fora.

A partir de 2003 o campeonato Brasileiro passou a ser disputado na forma de *round robin* e esse modelo permanece inalterado até os dias de hoje. Nesses anos, o campeonato brasileiro foi disputado entre os meses de maio a dezembro, após o encerramento dos campeonatos estaduais. Para maiores informações sobre o desenho do campeonato Brasileiro, ver Sant'Anna et al. (2007) e Sant'Anna et al. (2009) e consequente aproveitamento da tabela final do campeonato na classificação para outras competições, assim como o rebaixamento dos piores times. Este desenho despertou interesse por compreender melhor o *home advantage* na competição e as chances de cada time nos seguintes objetivos: ser classificado para uma das copas continentais como a copa Libertadores da América (ou seja, estar entre os 4 primeiros colocados), se classificar para a copa Sul-Americana (figurar entre o quinto e o décimo segundo lugar), ser campeão ou ser rebaixado (ou seja, figurar entre os 4 times de menor pontuação).

Nos campeonatos do tipo pontos corridos, a regularidade do time é considerada fundamental na sua classificação final. No campeonato Brasileiro com o formato de pontos corridos, cada par de times se enfrenta duas vezes, ou *round robin*, com cada partida sendo realizada com o mando de campo de cada time. Logo, a equipe necessita ser tão boa fora quanto dentro de casa. Sendo assim, o resultado de todos os jogos é de suma importância, pois o time mais eficaz ao longo de todo campeonato, será o campeão.

Outra forma de se realizar um campeonato é o eliminatório, como na copa do Brasil, onde as equipes se enfrentam em confronto direto e eliminatórios. Torna-se campeão o time que consegue eliminar todos os demais.

Uma mistura dos dois formatos supracitados é o campeonato no formato misto, onde o resultado da primeira fase serve como balizador para o confronto direto entre as equipes na fase seguinte. A equipe que tiver melhor desempenho na primeira fase terá algumas vantagens como decidir em casa e jogar por resultados iguais. Com isso uma vitória em casa por uma boa diferença de gols tem uma importância muito grande, na segunda fase.

Nesse formato de campeonato, pode-se destacar a copa do mundo e o campeonato Brasileiro até 2002. Os campeonatos que não possuem o formato de pontos corridos nem sempre premiam o “melhor time”, pois a continuidade do time na competição é dada exclusivamente por um confronto direto entre duas equipes. Isso porque existem vários fatores que podem influenciar o resultado desse confronto, como por exemplo: erros de

arbitragem e o momento do time na competição.

Como erro de arbitragem, vale citar o ocorrido na copa do mundo 2002, quando a Coréia do Sul acabou por eliminar a Itália e a Espanha em função de erros de arbitragem. Em relação ao melhor momento vivido pelo time, vale lembrar o campeonato brasileiro de 2002, onde a equipe do Santos sagrou-se campeã ao suplantar a equipe do Corinthians em dois jogos; ou mesmo na copa do Brasil de 2005, onde o time do Baraúnas eliminou a equipe do Vasco da Gama nas oitavas de final.

O objetivo do presente trabalho foi modelar o campeonato Brasileiro de 2008 objetivando acompanhar o desempenho de cada time ao longo do campeonato, considerando, como variáveis dependentes: a importância do mando de campo (*Home Advantage*), a força de cada time e o desempenho recente dos times. O objetivo é calcular chances de rebaixamento e classificação para a Copa Libertadores da América para os clubes.

O modelo proposto foi um modelo logito ordinal, também utilizado por Lawal et al (2007) na modelagem com campeonato Inglês (EPL) nos anos de 2002 e 2003.

Lawal et al (2007) trabalha com um modelo bastante simples, porém, eficaz, na previsão do EPL (*English Premier League*). O modelo utiliza a vantagem de jogar em casa e as forças individuais de cada time como variáveis preditivas.

Pollard (2002) analisa a redução da vantagem de uma equipe em jogar em casa quando esta se muda de uma cidade para outra, algo muito comum em esportes onde os times são franquias particulares ou em casos em que o time perde o mando de campo, por algum motivo. No futebol brasileiro existem casos mais particulares como os clássicos disputados na mesma cidade, como, por exemplo, os jogos entre as equipes do Rio de Janeiro, que são disputados no Maracanã. Essa situação não é exclusiva do Rio de Janeiro, ocorre também em Minas Gerais e São Paulo.

## 1.1 Objetivo

O objetivo principal do presente trabalho é avaliar a capacidade preditiva do modelo no que tange à classificação para a copa libertadores da América e o rebaixamento para a segunda divisão do campeonato brasileiro de futebol.

A variável resposta do modelo são as probabilidades para cada tipo de resultado de cada jogo (vitória, empate ou derrota do time mandante da partida) com base nos jogos previamente realizados.



Com base nessas probabilidades, serão realizadas simulações sobre os jogos não realizados para estimação da pontuação final de cada time. A partir desse resultado, calcula-se o número de vezes que um time alcança a copa libertadores da América (ficando entre os quatro times de maior pontuação) ou ser rebaixado para a segunda divisão do campeonato brasileiro (terminando entre os quatro times de menor pontuação).

Serão testados três modelos no presente trabalho, um deles considerando somente a força do time e a vantagem de jogar em casa; outros dois modelos agregam, de maneira diferente, o desempenho recente do time.

No primeiro modelo, todos os jogos tinham a mesma importância para o modelo, ou seja, existia uma inércia do modelo em perceber as mudanças que ocorriam ao longo do campeonato, em função disso, foram desenvolvidos os dois modelos seguintes, que tinham como objetivo eliminar essa inércia e agregar uma informação que pode ser considerada de suma importância no futebol, que é o momento da equipe, por isso que serão utilizadas variáveis que agregam os resultados de partidas mais recentes, disputadas pelas equipes.

## 1.2 Estrutura do Trabalho

Até o presente momento foram apresentados os temas, objetivos e a metodologia que será empregada no estudo. A partir de agora serão mostradas a disposição dos capítulos subsequentes.

O capítulo 2 abordará um breve histórico do campeonato brasileiro de futebol, assim como as mudanças ocorridas desde sua primeira edição.

O capítulo 3 será uma revisão bibliográfica sobre futebol, *home advantage* e previsão de resultados em esportes.

O capítulo 4 salientará a metodologia utilizada, que será o modelo logito ordinal, assim como técnicas de simulações para gerar as pontuações finais dos times.

No capítulo 5, o estudo explicitará os modelos trabalhados no presente trabalho.

No capítulo 6 traz a análise dos resultados obtidos.

Por fim, figurarão as conclusões do estudo e possíveis aplicações e melhorias futuras para o modelo.

## **2. CAMPEONATO BRASILEIRO DE FUTEBOL**

### 2.1 Evolução do Campeonato Brasileiro

O campeonato brasileiro de futebol teve início em 1971, entretanto, o mesmo já vinha sendo disputado desde 1959 com outros nomes. De 1959 a 1968, o campeonato era chamado de Taça Brasil; e, de 1967 a 1970, o campeonato era denominado Taça de prata. Nos anos de 1967 e 1968, foram disputados ambos os campeonatos concomitantemente.

O maior vencedor da Taça Brasil foi o Santos, que obteve cinco títulos dentre os dez disputados; pois, contava, em seu elenco, com jogadores como Pepe, Mengalvio e o melhor jogador de futebol de todos os tempos, Pelé. Além de ser o maior vencedor da Taça Brasil, o time do Santos foi bi-campeão mundial de futebol nesse período. A tabela 2.1 apresenta todos os campeões da Taça Brasil.

Tabela 2.1 – Campeões da Taça Brasil

<i>Times</i>	<i>Anos</i>
Santos	1961, 1962, 1963, 1964 e 1965
Palmeiras	1960 e 1967
Bahia	1959
Cruzeiro	1966
Botafogo	1968

Fonte: Revista do Campeonato Brasileiro 2009, Jornal O Globo

De 1967 a 1970 foram disputadas quatro edições da Taça de Prata, onde o Palmeiras conquistou duas edições dessa competição. A tabela 2.2 traz os campeões desta competição.

Tabela 2.2 – Campeões da Taça de Prata

<i>Times</i>	<i>Anos</i>
Palmeiras	1967 e 1969
Santos	1968
Fluminense	1970

Fonte: Revista do Campeonato Brasileiro 2009, Jornal O Globo

Vale ressaltar que a equipe do Palmeiras se sagrou campeã de ambas as competições (Taça Brasil e Taça de Prata) no ano de 1967.

A partir de 1971, a CBF (Confederação Brasileira de Futebol) passou a organizar o campeonato, denominando-o Campeonato Brasileiro de Futebol. A CBF organizou os campeonatos desde 1971 até 1986, porém, em 1987, a mesma não teve condições de organizar o campeonato da primeira divisão do campeonato brasileiro, passando essa tarefa ao Clube dos Treze (grupo formado pelos 13 maiores clubes do Brasil).

O Clube dos Treze foi fundado em 1987 com o intuito de defender os interesses políticos e comerciais dos seus membros. No início era composto por 13 clubes (Vasco da Gama, Corinthians, São Paulo, Palmeiras, Flamengo, Santos, Fluminense, Botafogo, Atlético Mineiro, Cruzeiro, Internacional, Grêmio e Bahia). Foi a partir de sua fundação que os clubes iniciaram campanhas de marketing com intuito de agregar valor à sua marca e tentar aumentar ganhos nesta área, pois até então os mesmos não exploravam suas marcas.

No ano de 1987, o campeonato foi chamado de Copa União e dividido em dois módulos principais: verde e amarelo. O módulo verde foi composto pelos times participantes do Clube dos Treze, que eram os clubes de maior notoriedade do país, além das equipes do Goiás, Santa Cruz e Coritiba e deixando de fora Guarani e América-RJ, que foram 2º e 4º lugares do campeonato de 1986.

Com toda a campanha de marketing e organização do campeonato em 1987, o mesmo teve impressionante média de público acima de 20 mil torcedores por jogo; fato este que não foi superado até hoje.

Na tabela 2.3 encontram-se os times que participaram dos módulos verde e amarelo da Copa União.

Tabela 2.3 – Times que Participaram da Copa União de 1987

<i>Módulo Verde</i>	<i>Módulo Amarelo</i>
Vasco da Gama	Guarani
Corinthians	Criciúma
Flamengo	Joinville
Fluminense	América-Rj
Botafogo	Internacional de Limeira
Grêmio	Portuguesa
São Paulo	Atlético-PR
Palmeiras	Rio Branco-ES
Santos	Bangu
Cruzeiro	Treze
Atlético-MG	Ceará
Bahia	CSA
Internacional	Náutico
Goiás	Atlético-GO
Santa Cruz	Sport
Coritiba	Vitória

Fonte: CBF

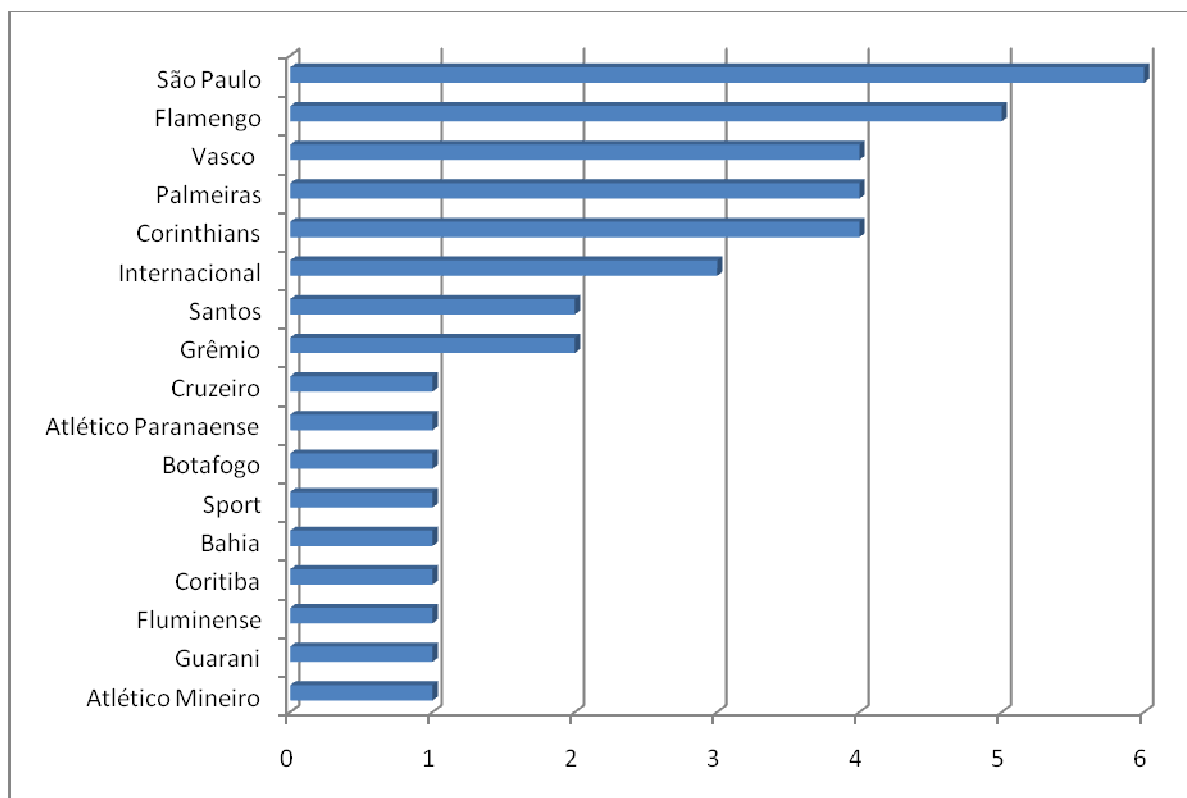
Contudo, o campeonato não transcorreu de forma natural; pois foi definido pelo conselho da federação no início do campeonato que os dois primeiros colocados de cada módulo (Verde e Amarelo) formariam um quadrangular para a definição do campeão brasileiro do ano. Contudo, não houve aceitação dos times que disputavam a competição no módulo verde, que se recusaram a jogar contra os dois primeiros colocados do módulo amarelo, por considerá-los inferiores. Em função disso, a CBF declarou Sport e Guarani (representantes do módulo amarelo) como campeão e vice da Copa União, respectivamente; em função da desistência de Internacional e Flamengo (representantes do módulo verde). Resultado que até hoje é contestado e bastante discutido

Vale ressaltar que, durante todos os 38 campeonatos realizados até o ano de 2008, apenas quatro times não tinham disputado a série B, que equivale a segunda divisão do campeonato brasileiro; que são: Flamengo, Internacional, Cruzeiro e Vasco da Gama. Porém, em 2009 a equipe do Vasco da Gama não disputa a série A do campeonato brasileiro, por ter

ficado entre os quatro times de menor pontuação no campeonato de 2008, sendo consequentemente rebaixado.

Vale ressaltar que, ao contrário de outros campeonatos realizados em outros países, existe uma alternância de campeões, isto é, em 38 edições do campeonato brasileiro, houve 17 campeões diferentes, onde, o maior campeão tem 6 títulos. O gráfico 2.1 ilustra o número de vezes que cada equipe foi campeã no campeonato brasileiro de futebol.

Figura 2.1 – Número de vezes que cada equipe foi campeã do campeonato brasileiro



Fonte: CBF

A equipe do São Paulo, além de ser o time com maior número de vitórias do campeonato brasileiro, também é o único tri-campeão consecutivo, conquistando os títulos de 2006, 2007 e 2008.

Essa grande diversidade de equipes com possibilidade de ganhar o título mostra o equilíbrio existente entre as equipes no campeonato brasileiro de futebol. Fato este que não ocorre em outras ligas, como pode ser observado na tabela 2.4:

Tabela 2.4 – Número de equipes que conquistaram o título em campeonatos pelo mundo

	Número de equipes que conquistaram o Título
Campeonato Brasileiro	<b>17</b>
Campeonato Inglês	10
Campeonato Alemão	9
Campeonato Francês	9
Campeonato Italiano	9
Campeonato Espanhol	7
Campeonato Português	5
Campeonato Holandês	4

Fonte: IFHHS – International Federation of Football History and Statistics

Para tornar possível a comparação entre os campeões, a tabela 2.4 ilustra o mesmo número de torneios disputados, desde o ano de 1971 até 2008, ou seja, 38 campeonatos.

Pode-se perceber que durante o período que o campeonato Brasileiro começou a ser disputado com este nome, o número de times que se sagraram campeões é muito maior que em ligas de grande expressão pelo mundo. Nestes países existe uma hegemonia por 3 ou 4 times, porém, em alguns períodos algum outro time de médio porte se destaca e acaba conquistando (o que não é tão comum), tanto que o campeonato que possui o maior número de campeões diferentes é o Inglês com apenas 10 times diferentes. Vale ressaltar que o número de equipes que disputam os campeonatos varia de 18 a 22, o que viabiliza a análise acima.

Alguns fatos relevantes na história do campeonato brasileiro de futebol:

- O Internacional é o único campeão brasileiro a terminar o campeonato sem derrota, em 1979;
- Em 1977, o São Paulo foi campeão ao vencer o Atlético/MG na decisão por pênaltis, após empate por 0 x 0. Nesse mesmo ano, tanto Atlético/MG, vice-campeão, como Botafogo, 5º colocado, terminaram o campeonato invictos;
- Um mesmo time nunca venceu a Taça Libertadores da América e o Campeonato Brasileiro no mesmo ano. Em 2006 o Internacional foi o primeiro a ser campeão da Libertadores e vice-campeão brasileiro na mesma temporada;

- Em 2003, o Cruzeiro foi o primeiro time a conquistar a "Tríplice Coroa" ao vencer no mesmo ano o título de seu estado (no caso o Campeonato Mineiro), a Copa do Brasil e o Campeonato Brasileiro;
- Por 12 vezes o Campeão Brasileiro também venceu o campeonato estadual no mesmo ano: Palmeiras (1972, 1993 e 1994), Internacional (1975 e 1976), Fluminense (1984), Bahia (1988), São Paulo (1991), Grêmio (1996), Corinthians (1999), Atlético Paranaense (2001) e Cruzeiro (2003);
- A equipe do Internacional foi a campeã com o melhor aproveitamento (em 1976) conseguiu 84,1% dos pontos disputados em 23 partidas. Já o Coritiba foi o campeão com o pior aproveitamento (em 1985), obtendo apenas 49,4% dos pontos disputados em 29 jogos.

Alguns jogadores fizeram história por se destacarem no campeonato brasileiro; seja por obter muitos títulos ou por sua capacidade técnica acima da média. Dentre os que conquistaram muitos títulos, destacam-se:

- Andrade, que conquistou 4 títulos pelo Flamengo e 1 pelo Vasco;
- Zinho, com 2 títulos pelo Flamengo, 2 pelo Palmeiras e 1 pelo Cruzeiro;
- Tita, com 3 títulos pelo Flamengo e 1 pelo Vasco;
- Junior, com 4 títulos pelo Flamengo.

Além desses, que se destacaram pela quantidade de títulos, existiram os jogadores cujo desempenho foi considerado muito acima do apresentado pelos demais jogadores da competição, culminando com sua transferência para equipes estrangeiras antes do término da competição, cerceando a oportunidade de ser campeão. Dentre esses jogadores, vale destacar: Zico, Sócrates, Junior, Romário, Robinho, Ronaldo, Ronaldinho Gaúcho, Kaká, Alexandre Pato, dentre outros.

Uma categoria pouco lembrada alguns anos atrás, porém, que vem ganhando cada vez mais destaque, são os técnicos de futebol. Em muitos casos, o ganho dos técnicos têm sido superior ao dos jogadores.

O técnico que foi mais vezes campeão do campeonato brasileiro até 2008 foi Vanderlei Luxemburgo, com 5 títulos, 1993 e 1994 pelo Palmeiras, 1998 pelo Corinthians, 2003 pelo Cruzeiro e 2004 pelo Santos. Além dele, destacam-se com três títulos cada, os técnicos:

- Rubens Minelli: Internacional (1975) e São Paulo (1977);
- Ênio Andrade: Internacional (1975), Grêmio (1981) e Coritiba (1985);
- Muricy Ramalho: São Paulo (2006, 2007 e 2008).

Três técnicos se destacaram como jogador e técnico: Emerson Leão, Muricy Ramalho e Paulo César Carpegiani. Emerson Leão chegou a exercer as duas funções concomitantemente em 1987, pois, estava inscrito como goleiro do Sport, porém, atuava como técnico, além do título de 1987, ele também conquistou o título de 2002, pelo Santos como técnico e mais 3 vezes como jogador (1972 e 1973 pelo Palmeiras e 1981 pelo Grêmio). Muricy Ramalho atuava como jogador de meio campo do São Paulo no ano de 1977, e é atual tri-campeão brasileiro pelo São Paulo (nos anos de 2006, 2007 e 2008). Paulo César Carpegiani foi três vezes campeão como jogador (em 1975 e 1976 pelo Internacional e em 1980 pelo Flamengo) e uma vez como técnico (1982 pelo Flamengo).

O primeiro gol na historia dos campeonatos brasileiros foi marcado por Hector Scotta em 7 de agosto de 1971 quando o mesmo atuava pelo Grêmio contra o São Paulo. A artilharia é um quesito que também é bastante disputado no campeonato brasileiro. Vale destacar o Vasco que já teve 8 vezes o artilheiro do campeonato. Além disso, o Vasco teve também o jogador Roberto Dinamite, o maior artilheiro de todos os tempos do campeonato brasileiro com 190 gols.

Atualmente o maior artilheiro dos campeonatos brasileiros é Washington, vulgo “Coração Valente”, em função de um grave problema no coração que quase fez com que o mesmo parasse de jogar; com 34 gols. Apesar de ter feito mais gols que o recordista anterior, Edmundo, Washington não conseguiu ultrapassar a impressionante média de mais de um gol por partida de Edmundo, que marcou 29 gols em 27 jogos.

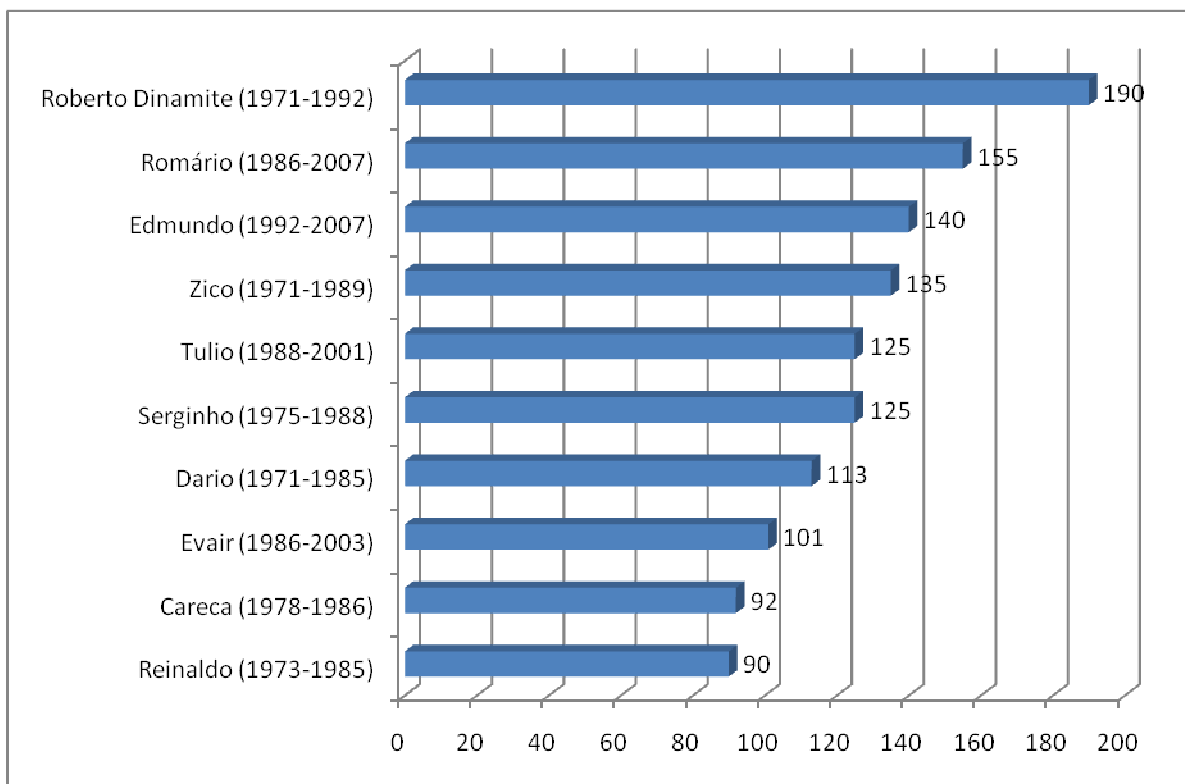
Durante a competição Edmundo conseguiu bater o recorde de número de gols marcados em uma única partida, marcando 6 gols contra o União São João, jogando pelo Vasco da Gama, no estádio do Vasco, em São Januário.

Túlio, Romário e Dário são os jogadores que mais vezes ganharam o título de artilheiro de um campeonato brasileiro, 3 vezes cada. Romário fez seu milésimo gol em um jogo do campeonato Brasileiro de 2007.

No gráfico 2.2 estão apresentados os 10 maiores artilheiros da história do campeonato brasileiro.



Figura 2.2 – Número de Gols dos Maiores Artilheiros da história do campeonato brasileiro



Fonte: CBF

O time que fez o maior número de gols em uma edição do campeonato brasileiro foi o Santos, com 103 gols no campeonato de 2004.

De 1971 a 2008 foram marcados 34.878 gols durante as 38 edições do campeonato em 14.472 jogos realizados, o que representa uma média de 2.41 gols por partida.

As 7 maiores goleadas estão apresentadas na tabela 2.4.

Tabela 2.4 – Maiores Goleadas do Campeonato Brasileiro até 2008

<i>Time Vitorioso</i>	<i>Placar</i>			<i>Time Derrotado</i>	<i>Data da Partida</i>
Corinthians	10	x	1	Tiradentes-PI	09/02/1983
Vasco	9	x	0	Tuna Luso-PA	19/02/1984
Flamengo	8	x	0	Fortaleza-CE	04/02/1981
Guarani	8	x	0	River –PI	03/02/1982
Flamengo	8	x	1	Sampaio Correa-PI	16/09/1976
Vitória	8	x	1	América-RN	23/03/1980
Guarani	8	x	1	Ceará	07/02/1982

Fonte: CBF

## 2.2 Formato de Disputa

O campeonato brasileiro teve diversos formatos desde a sua criação em 1971, assim como o número de participantes variou muito ao longo dos anos. Este tipo de formato atualmente é o mais utilizado no mundo, por isso, existem vários estudos que visam ajudar na elaboração das partidas, ou seja, como deve ser a organização da tabela do campeonato como em Urrutia et al. (2007) e Briskorn (2006). Nestes estudos são identificadas possíveis quebras no seqüenciamento das partidas e a melhor forma de organizar a tabela completa do campeonato.

Apenas em 2003 deu-se início a um processo de padronização dos campeonatos e uma adequação aos moldes dos campeonatos europeus; com a mudança no formato para pontos corridos.

Dentre os formatos utilizados no campeonato brasileiro, vale destacar o formato mata-mata, o de pontos corridos e os campeonatos mistos.

Os torneios no formato mata-mata são caracterizados pelo confronto direto dos participantes, com a eliminação do time derrotado; até que restem apenas dois competidores, responsáveis pelo último confronto e conseqüente definição do campeão. As copas nacionais, como, por exemplo, a Copa do Brasil, Copa da Itália e Copa da Inglaterra seguem esse formato de competição.

Tanto na Copa do Brasil quanto na da Itália, são disputados dois jogos entre cada par de equipes, onde cada equipe detém o mando de campo em uma das partidas; na copa da

Inglaterra, existe um sorteio prévio para a definição do mandante de campo e é realizado apenas um jogo.

Este formato da Copa da Inglaterra muitas vezes favorece a times pequenos, que acabam conseguindo façanhas como eliminar grandes forças do futebol mundial. Duas surpresas recentes foram o Guingamp (que é considerado um time mediano da segunda divisão), atual campeão da Copa da França. Outras surpresas podem ser consideradas os títulos de Portsmouth (Inglaterra), Nürnberg (Alemanha), Roda e Utrecht (ambos da Holanda) e Vitória de Setúbal (Portugal).

A abrangência das copas nacionais é sempre muito maior do que o campeonato nacional, pois, qualquer time de qualquer divisão pode participar; inclusive times amadores. A única exceção ocorre na Copa do Brasil, onde os participantes são selecionados a partir de seu desempenho nos campeonatos estaduais do ano anterior. Em função disso, as surpresas supracitadas ocorrem em menor grau.

Nos campeonatos do tipo pontos corridos, ou *round robin*, a regularidade do time é considerada fundamental na sua classificação final. Nesse formato, cada equipe enfrenta todas as demais duas vezes, uma em seus domínios e outra no estádio definido pelo adversário. Ao final, o time de maior pontuação se sagra campeão. Esse é o formato dos principais campeonatos europeus e, desde 2003, é também o formato do campeonato brasileiro.

Nesse tipo de formato, a equipe necessita ser tão boa fora, quanto dentro de casa. Sendo assim, o resultado de todos os jogos é de suma importância, pois o time mais eficiente ao longo de todo campeonato será o campeão.

Enquanto isso, os campeonatos mistos se caracterizam por apresentar os dois formatos acima citados. Isso é feito através de uma divisão em duas fases onde, a primeira fase é uma divisão dos times em grupos, onde todos jogam contra todos dentro de cada grupo. Essa primeira fase serve basicamente para classificar os times para uma segunda fase, que será disputada no formato mata-mata.

A equipe que tiver melhor desempenho na primeira fase terá algumas vantagens como decidir em casa e jogar por resultados iguais. Com isso uma vitória em casa por uma boa diferença de gols tem uma importância muito grande.

Como exemplo desse formato de campeonato, vale destacar a Copa do Mundo, a Copa Libertadores da América e a Copa dos Campeões da Europa.

No futebol, o momento vivenciado por cada time ou atleta é muito valorizado, tanto que existe um clichê “futebol é momento”. Tanto que times tidos como excelentes podem perder para times considerados com um menor nível técnico. Isto foi evidenciado no campeonato de 2002, onde o São Paulo possuía a melhor campanha durante toda a competição e acabou sendo eliminado pelo Santos, que havia se classificado nos últimos minutos da última rodada. Outros exemplos são Santo André e Paulista de Jundiaí, que se sagraram campeões da Copa do Brasil ao derrotar Flamengo e Fluminense, respectivamente, que são times considerados de nível técnico superior aos times citados anteriormente.

### 2.3 Mudanças de Regra e Formato dos Campeonatos Brasileiros

O Campeonato Brasileiro que se iniciou em 1971 se destaca principalmente pela mudança constante de regras assim como por mudanças no número de participantes. Durante as primeiras 10 edições apenas nos anos de 1973 e 1974 que o número de participantes permaneceu o mesmo. Esta constante mudança de formatos e regras para a disputa de um campeonato atrapalha o entendimento do campeonato. Haigh (2009) cita que as regras devem ser mantidas para facilitar a participação e criar uma maior familiaridade das pessoas com as regras que estão sendo utilizadas.

Na tabela 2.5 são apresentadas todas as edições de campeonato e o número de participantes de cada uma.

Tabela 2.5 – Número de Equipes Disputando o Campeonato até 2008

<i>Ano</i>	<i>Numero de Times</i>
1971	20
1972	26
1973 e 1974	40
1975	42
1976	54
1977	62
1978	74
1979	94
1980 até 1983	44
1984 e 1985	41
1986	48
1987	16 em cada módulo
1988	24
1989	22
1990 até 1992	20
1993	32
1994 até 1996	24
1997	26
1998	24
1999	22
2000	29 / 116
2001	28
2002	26
2003 e 2004	24
2005	22
2006 até 2008	20

Fonte: CBF

Não só de aumento e diminuição de participantes vivia o campeonato brasileiro, pois em todos os anos acontecia uma mudança de formato. Por exemplo, em 1971 o campeonato foi disputado com 20 participantes divididos em 2 grupos de 10 times e todos jogavam contra todos, classificando os 6 primeiros times de cada grupo. Essas 12 equipes seriam posteriormente divididas em 3 grupos de 4, onde os 3 primeiros colocados de cada grupo decidiriam o campeonato em turno único.

Diversas mudanças de formato foram sucedendo ano após ano.

Em 1976 o campeonato foi disputado por 54 clubes divididos em 6 chaves, onde se classificavam os 4 melhores de cada chave. Os 5 restantes formavam uma repescagem através de 6 chaves de 5 clubes cada, onde o campeão de cada chave garantia o direito de participar da segunda fase com os demais 4 previamente classificados. Nesta segunda fase eram

formadas 4 chaves de 6 clubes cada; e os 3 melhores de cada chave se classificavam para a terceira a fase.

Os 18 classificados (12 oriundo da segunda fase e 6 da repescagem) eram divididos em 2 grupos de 9 clubes que se enfrentavam em turno único dentro de cada grupo e se classificavam os dois primeiros de cada grupo para as semi-finais. O enfrentamento é baseado na chave olímpica, onde o primeiro colocado de um grupo enfrenta o segundo colocado do outro grupo e vice-versa em um jogo único e com mando de campo para os primeiros colocados. Por fim, é realizada a final em jogo único com mando de campo favorecendo o time de melhor campanha. Em caso de empate, o jogo vai para prorrogação e caso o empate persista o campeão será decidido nos pênaltis.

A criatividade dos dirigentes brasileiros não tinha limites e no ano seguinte o regulamento foi tão incomum quanto de 1976. Pode-se destacar neste ano o agrupamento das equipes, onde cada grupo tinha um número diferente de times.

Quando se pensava que o número de times pararia de crescer os dirigentes inseriam mais clubes, com o objetivo de se favorecer em eleições para confederações estaduais. Em 1978 o número de equipes suplantou o de 1977, atingindo o incrível número de 94 equipes.

Os regulamentos se tornavam cada vez mais complexos e com mais equipes, porém na década de 80 iniciou-se um enxugamento do número de clubes, pois via-se que os campeonatos ganhavam cada vez mais fases e os times acabavam jogando mais, sem que tivesse interesse nos jogos iniciais. Algumas regras merecem ser destacadas, como por exemplo:

- Nos campeonatos de 1975 até 1977 eram concedidos 1 ponto a mais para as equipes que vencessem por 2 ou mais gols de diferença;
- Em 1978 a regra se manteve, porém apenas para as vitórias por 3 ou mais gols de diferença.

Conforme os anos avançavam o formato do campeonato ia se adaptando e modificando em função da redução do número de times, o que de certa forma melhorou e facilitou o entendimento das regras e regulamentos.

Em 1987 houve problemas na hora de organizar o campeonato e a CBF não chegou a um consenso com os clubes e estes fundaram o Clube dos Treze, que ajudou a organizar o campeonato daquele ano. Porém, houve desentendimento entre as partes e o campeonato acabou sendo decidido na justiça, o que é motivo de muita discussão até hoje.

Em 1988 foi inserida a disputa de pênaltis para todos os jogos que terminassem empatados, para isso foi modificada a pontuação do campeonato, onde o vencedor recebia 3 pontos pela vitória no tempo normal, enquanto que o vencedor na disputa por pênaltis recebia 2 pontos e a equipe derrotada recebia 1 ponto.

De 1989 a 1994 os campeonatos passaram a ter a pontuação definida pelas regras da FIFA e as disputas de pênalti foram abolidas.

Em 1995, a FIFA, com o intuito de aumentar a motivação das equipes para tentar conseguir a vitória, determinou que cada vitória passasse a valer 3 pontos, o que modificou a pontuação no campeonato brasileiro. **## Acrescentar EJOR partoni**

Na década de 90 o regulamento do campeonato não sofreu tantas modificações de ano para ano, pois o formato básico era o confronto direto entre as equipes, onde os 8 primeiros classificados eram divididos em 2 grupos de 4 e o melhor de cada grupo se classificava para a final, que era disputada em 2 jogos.

Em 2000, em função de um problema enfrentado pela CBF com o caso da alteração de idade de um jogador no campeonato do ano anterior, a CBF ficou impossibilitada de organizar o campeonato, pois teve de esperar que o caso se resolvesse. Em função disso, o Clube dos Treze assumiu e organizou o campeonato que se chamou Copa João Havelange e voltou a contar com mais de 100 times, divididos em quatro módulos: azul, amarelo, verde e branco.

O módulo azul era composto por 25 times que teriam o direito de disputa a série A do campeonato brasileiro, além do Gama, que ganhou o direito de disputar na justiça em função do uso do jogador Sandro Hiroshi pelo São Paulo, logo o Gama ganhou os pontos dessa partida e não seria rebaixado. O módulo Amarelo correspondia aos 36 times da série B. Os módulos branco e verde correspondiam a série C dividida por região, para que os custos de viagem fossem diluídos, desta forma, agrupou-se as Regiões Sul e Sudeste no módulo Branco e as demais regiões no módulo Verde.

Vale ressaltar que na final houve problemas no jogo disputado em São Januário e o campeão só foi consagrado em 2001.

A partir de 2003 o campeonato brasileiro passou a ser disputado em ponto corridos. Iniciou-se a competição por pontos corridos com 24 times e o número de participantes foi diminuindo de forma gradativa com o descenso de 4 equipe e o acesso de apenas 2, o que de certa forma dificultava o acesso a primeira divisão. A partir de 2006, esse formato se estabilizou com apenas 20 equipes, onde passaram a subir 4 equipes e a descer outras 4.

Esta forma pode ser considerada mais justa, pois acaba por se sagrar campeão o time mais regular ao longo da temporada e não apenas o time que vive um momento melhor na hora certa.

Um ponto que ainda precisar ser melhorado no futebol brasileiro é a adequação ao calendário europeu, pois existe a janela de meio de ano, onde os times europeus têm o mercado aberto e acaba por levar muitos jogadores neste período, o que acaba por enfraquecer os principais times brasileiros. Este problema pode ser visto de forma bastante clara em 2007, quando os times brasileiros perderam seus principais jogadores no meio do campeonato, ocasionando um pior desempenho das equipes no segundo turno.



### **3. REVISÃO BIBLIOGRÁFICA**

Nos últimos anos, diversos trabalhos envolvendo esportes foram desenvolvidos. Muitos têm como foco a avaliação da existência do *home advantage*, ou seja, vantagem de jogar em casa em vários esportes. Essa vantagem impacta no resultado final dos jogos e é defendida por muitos autores, sendo aplicada tanto no futebol como no beisebol, voleibol, futebol americano, entre outros esportes (Courneya e Carron, 1992; Pollard, 2005).

Ainda relacionado ao *home advantage*, Pollard (2006) realizou um estudo com 6 temporadas de jogos realizados em ligas regionais de 72 países e identificou o *home advantage* como principal variável explicativa

Dare et. al (1996) fazem uso de um modelo linear generalizado mesclando informações do mercado de apostas com *home advantage*. O autor avalia a relação da equipe tida como favorita pelo mercado de apostas; e pelo time que possui a vantagem de jogar em casa (caso o campo não seja neutro).

O futebol é um esporte que possui grande penetração em todo o planeta, por se tratar de um esporte que não exige grandes materiais para sua execução. Desta forma, este é assunto de grande interesse por parte de diversos autores.

Nevill et. al (1996) fizeram um estudo na liga inglesa e escocesa de futebol para os anos de 1992-1993 e identificaram que a torcida tem um impacto no resultado da partida. O autor faz uma comparação entre as divisões do campeonato, constatando que divisões com maior público, o time da casa obtinha mais vantagem que o time visitante. O autor tem duas explicações para isso; o peso da torcida fazia com que os adversários errassem mais os pênaltis, assim como o árbitro tendia a achar que o time visitante cometia mais infrações que os times da casa.

Times que possuem uma história de jogos internacionais ou que já participaram de muitos campeonatos, possuem mais *expertise* e habilidades que farão com que sua *performance* seja melhorada. Gelade e Dobson (2007) fizeram um estudo comparativo entre dos times de 204 países *rankeados* pela FIFA entre os anos de 2000 a 2005 e identificaram que a tradição do time influenciava na *performance* do mesmo.

Estudos relacionados à previsão em esportes são bastante comuns, principalmente porque países como Inglaterra, Irlanda, Alemanha e outros, possuem casas de apostas regularizadas, que movimentam muito dinheiro, gerando maior interesse neste tipo de estudo.

Spann e Skiera (2008) tentam prever os resultados dos jogos do campeonato Alemão de futebol através de diversas técnicas, a partir dos dados das previsões, tentam estabelecer regras sobre o desempenho de cada equipe. McHale e Scarf (2006) utilizaram jogos realizados no período de 1993 a 2004 do campeonato inglês para desenvolver um modelo que conseguisse prever de forma adequada o resultado de jogos de futebol, para isso eles tentam montar os modelos através das distribuições marginais e a dependência de variáveis, o que ajuda a definir o resultado final das partidas.

Cheng et al. (2003) utiliza redes neurais híbridas para calcular a probabilidade de um determinado time A bater um outro time B. Isto foi feito através do uso de redes neurais, onde os times foram classificados conforme sua força; a partir dessa classificação, a rede neural é treinada e testada com os resultados reais e por fim alguns jogos são previstos e por fim comparados aos resultados reais. Vale ressaltar que o uso de redes neurais para prever resultados de jogos não é algo muito comum no meio acadêmico.

Yiannakis et al. (2006) utilizou uma série temporal, com base no modelo ARIMA, com o intuito de determinar o quanto os resultados recentes interferem no desempenho do time nas partidas subsequentes, para isso, o autor fez uso dos últimos 10 jogos de ambas as equipes.

Além de trabalhos que têm metodologias por trás para embasar uma análise de futebol, vale destacar alguns qualitativos, como em Andersen et al. (2005), que fez uma análise junto a 251 pessoas, com o objetivo de prever as equipes classificadas para a segunda fase da Copa do Mundo de 2002.

Outra área com muitas aplicações em esportes é DEA (*Data Envelopment Analysis*) com criações ou verificações de *ranking*. Churilov and Flitman (2006) propõem uma nova forma de *rankear* os países participantes das olimpíadas de verão com base nas medalhas conquistadas e a população do país, através de um modelo DEA. Lins et al. (2003) criticam a

maneira que o COI (Comitê Olímpico Internacional) utiliza para montar o *ranking* dos países nos jogos olímpicos de verão e propõe uma abordagem DEA para estabelecer o *ranking*, utilizando as medalhas e o produto interno bruto dos países.

Calôba e Lins (2006) fazem uma comparação dos *rankings* da CBF e da revista Placar (revista nacional de grande público, focada em futebol, que faz um ranking com base em uma pontuação diferente da estabelecida pela CBF).

Sant'Anna et al. (2009) disserta sobre o formato de disputa do campeonato brasileiro e conseqüente aproveitamento da tabela final do campeonato na classificação para outras competições, assim como o rebaixamento dos piores times.

## **4. METODOLOGIA**

A partir deste capítulo serão abordados os temas relativos à metodologia utilizada no presente estudo. Algumas das metodologias que serão utilizadas e explicitadas a partir de agora são: Modelos lineares generalizados, Regressão Logística e Simulação de Resultados.

### **4.1- Modelos Lineares Generalizados (MLG)**

Nelder e Weddeburn (1972) desenvolveram uma classe de modelos baseados na família exponencial com um parâmetro desconhecido, cujas médias são não-lineares num conjunto de parâmetros lineares. Tais modelos são chamados de modelos lineares generalizados e tem aplicações em diversas áreas de interesse, como por exemplo:

- Esportes;
- Área medica;
- Transportes;

Dentre os diversos métodos que são classificados como modelos lineares generalizados e que têm mais aplicações estão:

- Modelo clássico de regressão com erro normal;
- Modelo log-linear aplicado a análise de tabelas de contingência;
- Modelo logístico para tabelas multidimensionais de proporções;
- Modelo “probit” para estudo de proporções;
- Modelo de análise de variância com efeitos aleatórios;
- Modelo estrutural para dados com distribuição gama;
- Análise de regressão não-simétrica;

- Polinômios inversos;
- Modelos de testes de vida.

Modelos lineares generalizados têm sua importância e grande aplicabilidade, pois possuem 3 etapas bem definidas como a formulação e ajustamento do modelo e por fim a inferência. Estes modelos acabam sendo de grande praticidade, por apresentar grande flexibilidade na sua formulação, computação simples de fazer para o ajustamento do modelo e o uso de critérios bastante razoáveis e fundamentados para a parte inferencial.

Para a parte da formulação do modelo é necessário definir toda a informação que será utilizada no modelo; é neste momento que são definidas as informações mais importantes para o ajustamento e a estimação dos resultados como a distribuição de probabilidade da variável resposta, co-variáveis e as funções de ligação.

A definição da distribuição de probabilidade da variável resposta exige do analista grande conhecimento e cuidado, pois faz-se necessário examinar alguns pontos cuidadosamente, como: assimetria e natureza dos dados e o intervalo de variação. Ao avaliar estes pontos, o analista conseguirá definir qual distribuição se aproxima melhor dos dados coletados.

Assim como a distribuição de probabilidade da variável resposta, outros dois pontos exigem muito do conhecimento e da análise do responsável pelo projeto, que é a definição das co-variáveis de forma eficiente, pois com isto a explicação do modelo pode ser melhorada.

Em relação à função de ligação, é necessário um estudo bastante abrangente sobre a distribuição dos erros. Para isso, faz-se necessária uma boa definição a priori, um exame intensivo dos dados e uma boa interpretação do modelo.

Para o modelo linear clássico, a identidade é utilizada como função de ligação, pois os valores esperados dos dados e das variáveis preditivas podem assumir qualquer valor. Quando o tipo de dado é uma contagem, utiliza-se uma Poisson.

Na tabela 4.1 são apresentadas o cruzamento entre a distribuição da variável resposta e função de ligação. O uso de função de ligação visa ter um bom ajuste para os seguintes efeitos: simetria do logaritmo da verossimilhança, constância da informação de Fisher e curvatura do logaritmo da verossimilhança.

Tabela 4.1: Cruzamento entre a Distribuição da variável resposta e da Função de Ligação

Distribuição da Função de Ligação	Distribuição da Função da Variável Resposta				
	Normal	Poison	Binomial	Gama	Normal Inversa
Identidade	X			X	
Logaritmo		X			
Inversa	X			X	
Inverso do Quadrado					X
Logística			X		
“Probit”			X		

Fonte: Modelos Lineares Generalizados. Cordeiro (UFPE)

A escolha de co-variáveis é importante, pois, definir este conjunto é essencial para representar a estrutura linear do modelo linear generalizado, pois assim forma-se a matriz modelo. Estas co-variáveis podem ser contínuas, qualitativas e mistas.

Dependendo do número de co-variáveis existentes, pode-se utilizar a inclusão da interação dos fatores. Entretanto, a interação é algo que nem sempre pode ser utilizado e quando feito precisa ser feito com bastante cuidado e parcimônia.

Após a execução dessa primeira fase, pode-se iniciar o ajustamento do modelo, com a estimação dos parâmetros lineares dos modelos e de determinadas funções destes parâmetros, que representam medidas de adequação dos valores estimados.

Para a estimação dos parâmetros existem várias opções como os estimadores de máxima verossimilhança, estimação por mínimos quadrados generalizados e outros. O método mais utilizado é o de máxima verossimilhança.

#### 4.1.1- Modelo de Regressão Logística Binária

Em modelos lineares generalizados existe um modelo onde as variáveis dependentes do modelo apresentam valores dicotômicos, ou seja, valores com duas categorias apenas, a classe de modelos que resolve esses problemas são os chamados de regressão logística binária, por tratar apenas de casos onde a resposta que se tem é uma variável binária.

Para tal, o tratamento logístico para a variável esta na função 4.1, que é a função onde o cálculo do modelo logístico se baseia:

$$Y = \frac{e^{\beta_0 + \sum_{j=1}^{p-1} \beta_j X_j}}{1 + e^{\beta_0 + \sum_{j=1}^{p-1} \beta_j X_j}} \quad (4.1)$$

Onde:

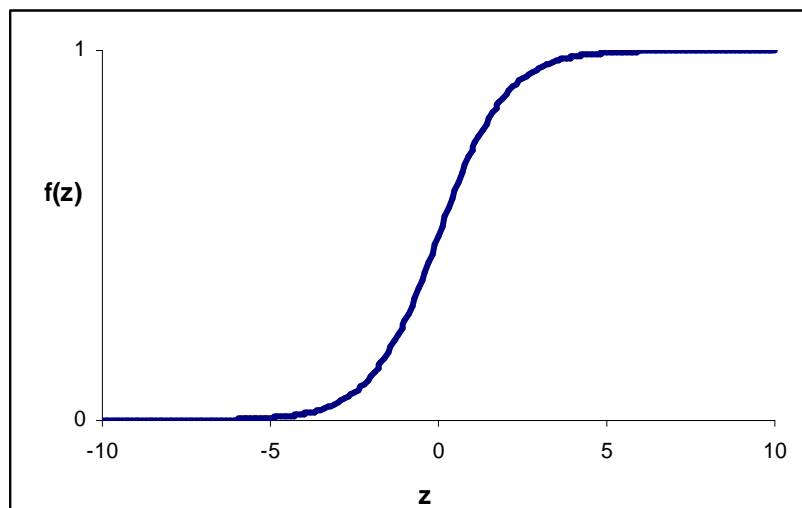
$Y \rightarrow$  variável dependente do modelo;

$\beta \rightarrow$  vetor de coeficientes do modelo;

$X \rightarrow$  vetor de variáveis independentes.

O formato da função sigmóide que faz a ligação entre a variável resposta e o modelo está ilustrado na figura 4.1.

Figura 4.1 – Função Sigmóide



Ao observar o gráfico, pode-se perceber que conforme o crescimento da variável independente a variável dependente se aproxima de 1 e conforme ela decresce a variável  $Y$  se aproxima de zero. Em função destas características que o modelo binário logístico é bastante difundido, pois ele serve para descrever a probabilidade de um evento.

Para aplicar o modelo, faz-se necessário linearizar a função logística a partir da transformação descrita em 4.2:

$$\text{logit}P(x) = \ln\left(\frac{P(x)}{1 - P(x)}\right) \quad (4.2)$$

Substituindo  $Y$  descrito em (4.1) por  $P(x)$  descrito em (4.2), tem-se:

$$\log itP(x) = \ln \left( \frac{\frac{e^\mu}{1+e^\mu}}{1 - \frac{e^\mu}{1+e^\mu}} \right)$$

Onde  $\mu = \beta_0 + \sum_{j=1}^{p-1} \beta_j X_j$

Resolvendo a equação acima, obtém-se:

$$\log itP(x) = \ln \left( \frac{\frac{e^\mu}{1+e^\mu}}{\frac{1}{1+e^\mu}} \right)$$

$$\log itP(x) = \ln(e^\mu)$$

Por fim chega-se ao final da expressão:

$$\log itP(x) = \mu$$

$$\log itP(x) = \beta_0 + \sum_{j=1}^{p-1} \beta_j X_j \quad (4.3)$$

A expressão 4.3 é a regressão logística linear, ou seja, após a transformação (4.2), o modelo logit passa a respeitar e pode utilizar todas as propriedades decorrentes do modelo de regressão linear, pois o logit passa a ser linear em todos os parâmetros.

#### 4.1.2- Modelo Logístico Simples

Nesta parte da metodologia, será apresentada a forma mais comum do modelo logístico, onde a amostra possui n pares de observações X e Y. Vale ressaltar que a variável dependente será uma variável dicotômica, ou seja, zero ou um e outra variável explicativa X. Com isso o modelo será definido da seguinte forma:

$$P(x) = \frac{e^{\beta_0 + \beta_1 X_1}}{1 + e^{\beta_0 + \beta_1 X_1}}$$

Onde:

Y → variável dependente do modelo;

$\beta_0$  → coeficiente linear do modelo.

$\beta_1$  → coeficiente angular do modelo.



A estimação dos parâmetros é feita através da máxima verossimilhança, ou seja, o objetivo é maximizar a função de verossimilhança. Após a estimação dos parâmetros serão encontradas as probabilidades para que a variável dependente  $Y$  seja igual a 1, para um dado valor de  $X$ . Logo a  $P(Y=1/X)=P(x)$ .

Em contrapartida a probabilidade do evento a ser estudado ser zero é o complementar da probabilidade supracitada. Logo, tem-se  $P(Y=0/X)=1-P(x)$ .

Ao apresentar tais características, pode-se dizer que a variável estudada apresenta uma distribuição de probabilidade Bernoulli, ou seja,  $Y \sim \text{Bernoulli}(P(x))$ , com a seguinte função de distribuição:

$$f(Y_i / X_i) = [P(x_i)]^{Y_i} [1 - P(x_i)]^{1 - Y_i}$$

Agora serão demonstrados e calculados os estimadores dos parâmetros a partir do método de máxima verossimilhança; para isso, faz-se necessário definir a função de verossimilhança, da seguinte forma:

$$\lambda(\beta) = \prod_{i=1}^n f(Y_i / X_i)$$

Os valores estimados para os parâmetros desta função são aqueles que fazem com que a função atinja o valor máximo. Diante disto, para que o tratamento das equações seja facilitado, é melhor que toda a parte instrumental seja calculada após o uso da função logarítmica na função:

$$L(\beta) = \ln[\lambda(\beta)]$$

$$L(\beta) = \ln \prod_{i=1}^n f(Y_i / X_i)$$

$$L(\beta) = \sum_{i=1}^n \{Y_i \ln[P(X_i)] + (1 - Y_i) \ln[1 - P(X_i)]\}$$

$$L(\beta) = \sum_{i=1}^n \left\{ Y_i \ln \frac{e^{\beta_0 + \beta_1 X_i}}{1 + e^{\beta_0 + \beta_1 X_i}} + (1 - Y_i) \ln \left[ 1 - \frac{e^{\beta_0 + \beta_1 X_i}}{1 + e^{\beta_0 + \beta_1 X_i}} \right] \right\}$$

$$L(\beta) = \sum_{i=1}^n \left\{ Y_i (\beta_0 + \beta_1 X_i) - \ln [1 + e^{\beta_0 + \beta_1 X_i}] \right\}$$

Para que o valor máximo da função seja encontrado é necessário que se tire a primeira derivada da função, e em seguida igualá-la a zero, com isto serão encontrados os valores dos parâmetros que maximizam a função. A derivada da função e, função de  $\beta_0$  e  $\beta_1$  está

apresentada em (4.4) e (4.5). Isso é válido quando a segunda derivada da função for menor do que zero.

$$\frac{\partial L(\beta)}{\partial \beta_0} = \sum_{i=1}^n \left\{ Y_i - \frac{e^{\beta_0 + \beta_1 X_i}}{1 + e^{\beta_0 + \beta_1 X_i}} \right\} \quad (4.4)$$

$$\frac{\partial L(\beta)}{\partial \beta_1} = \sum_{i=1}^n \left\{ X_i Y_i - \frac{X_i e^{\beta_0 + \beta_1 X_i}}{1 + e^{\beta_0 + \beta_1 X_i}} \right\} \quad (4.5)$$

Outra forma de apresentar esta formulação seria a seguinte:

$$\frac{\partial L(\beta)}{\partial \beta_0} = \sum_{i=1}^n \{ Y_i - P(X_i) \}$$

$$\frac{\partial L(\beta)}{\partial \beta_1} = \sum_{i=1}^n X_i \{ Y_i - P(X_i) \}$$

Um problema dos modelos de regressão logística é que as expressões a serem calculadas são não lineares, o que demanda a utilização de métodos iterativos para a resolução do problema. Estes métodos definem um valor inicial para os parâmetros e vão sendo calculados os valores para os parâmetros de maneira iterativa, quando a diferença entre o valor do parâmetro atual e do parâmetro chega a um determinado valor definido pelo analista, as iterações param e os parâmetros utilizados serão estes.

#### 4.1.3- Modelo Logístico Múltiplo

A partir de agora haverá uma mudança no número de variáveis independentes, pois neste momento será utilizado um vetor de dimensão 1-p de variáveis independentes, que pode ser expresso por  $X' = [X_1, X_2, X_3, X_4, \dots, X_{1-p}]$ . Diante das observações das 1-p variáveis dependentes o modelo de regressão logística será escrito na forma:

$$Y = \frac{e^{\beta_0 + \sum_{j=1}^{p-1} \beta_j X_j}}{1 + e^{\beta_0 + \sum_{j=1}^{p-1} \beta_j X_j}}$$

Onde:

$Y \rightarrow$  variável dependente do modelo;

$\beta \rightarrow$  vetor de variáveis independentes do modelo.

Assim como no modelo univariado, a estimação dos parâmetros será feita utilizando os estimadores de máxima verossimilhança, logo, todos os passos feitos na parte anterior seriam aplicáveis agora, para tanto, serão ilustrados somente os passos a partir da derivada.

$$\frac{\partial L(\beta)}{\partial \beta_0} = \sum_{i=1}^n \{Y_i - P(X_i)\} = 0$$

$$\frac{\partial L(\beta)}{\partial \beta_j} = \sum_{i=1}^n X_{ij} \{Y_i - P(X_i)\} = 0 \text{ para } j=0,1,2,\dots,p-1$$

Como no modelo univariado, tem-se equações não lineares, porém será encontrado um vetor de soluções de tamanho  $p-1$  com todos os estimadores  $\hat{\beta}$ .

Outra possibilidade para o cálculo dos estimadores seria a utilização da minimização da soma dos quadrados dos resíduos, com a seguinte formulação:

$$\min \sum_{i=1}^n \left[ y_i - \hat{P}(X_i) \right]^2 = \min SQR$$

Segundo Walker e Duncan (1967) apud Guimarães (2000) e Martins (1988) encontrar de forma analítica o valor dos parâmetros é algo bastante difícil. Em função disso alguns métodos numéricos, como Newton-Raphson, são utilizados para que a convergência dos valores seja feita de forma mais rápida.

Para a determinação dos estimadores pode-se usar os dois métodos citados anteriormente, porém no momento de definirmos as variâncias e covariâncias das variáveis é necessário um trabalho um pouco maior, pois apesar do teorema do limite central dizer que os estimadores são assintoticamente normais, é necessário calcular segunda derivada parcial da função de log verossimilhança, como descrito em (4.6) e (4.7):

$$\frac{\partial^2 L(\beta)}{\partial \beta_j^2} = - \sum_{i=1}^n X_{ij}^2 P(X_i) \{1 - P(X_i)\} \quad (4.6)$$

$$\frac{\partial^2 L(\beta)}{\partial \beta_j \partial \beta_u} = - \sum_{i=1}^n X_{ij} X_{iu} P(X_i) \{1 - P(X_i)\} \quad (4.7)$$

Para  $j,u=0,1,2,\dots,p-1$ . e  $i=1,\dots,n$

A construção da matriz de variância e covariância demanda tempo, principalmente por se tratar de uma matriz de dimensão  $p$ . Para isso serão utilizados os termos das equações (4.6) e (4.7), além de fazer uso de uma matriz identidade para o número de estimadores  $\beta$ , ou seja, uma matriz  $I(\beta)$ , que é a matriz de informação. Logo a variância e covariância serão obtidas através da inversão desta matriz de informação, denotado por  $\Sigma(\beta) = I^{-1}(\beta)$ .

Para simplificar a notação e facilitar a vida do leitor, foi estabelecida uma regra bastante utilizada em toda literatura para identificar as variâncias e covariâncias de cada elemento na matriz. O  $j$ -ésimo elemento da diagonal principal da matriz será a variância do

parâmetro  $\hat{\beta}_j$  e será denotado na matriz como  $\sigma^2(\hat{\beta}_j)$  e será utilizado  $\sigma(\hat{\beta}_j, \hat{\beta}_u)$  para

denotar a covariância entre os parâmetros  $\hat{\beta}_j$  e  $\hat{\beta}_u$ , onde  $u \neq j$  e  $j, u = 0, 1, 2, \dots, p-1$ .

Por fim chega-se que os desvios padrões das estimativas utilizadas serão explicitados na equação 4.8:

$$\sigma(\hat{\beta}_j) = \left[ \sigma^2(\hat{\beta}_j) \right]^{1/2} \quad \text{para } j=0,1,2,\dots,p-1. \quad (4.8)$$

O uso dos desvios padrões para as estimativas são bastante utilizados para testar a hipótese da nulidade dos coeficientes e para a estimação de intervalos de confiança para as estimativas encontradas.

Diante do que foi exposto, agora será apresentada a formulação para a matriz de informação que é:

$$I(\hat{\beta}) = X' V X$$

Onde:

$X \rightarrow$  é uma matriz de  $n \times p$  elementos, que possui toda a informação de cada elemento

$V \rightarrow$  é a matriz diagonal de dimensão  $n \times n$  com elemento geral  $\hat{P}(X_i) \left[ 1 - \hat{P}(X_i) \right]$ .

Logo, as matrizes citadas possuem o seguinte formato:

$$X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1(p-1)} \\ 1 & x_{21} & \dots & x_{2(p-1)} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & \dots & x_{n(p-1)} \end{bmatrix} e$$

$$V = \begin{bmatrix} \hat{P}(X_1)[1-\hat{P}(X_1)] & 0 & \dots & 0 \\ 0 & \hat{P}(X_2)[1-\hat{P}(X_2)] & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & \hat{P}(X_n)[1-\hat{P}(X_n)] \end{bmatrix}$$

Para finalizar, será mostrado o cálculo da função *score* para o método de Newton-Raphson, que utiliza a formulação (4.9) para efetuar o cálculo dos mínimos quadrados dos resíduos:

$$\beta_{k+1} = \beta_k + \{I(\beta_k)\}^{-1} \{s(\beta_k)\} \quad (4.9)$$

Onde

$$s(\beta_k) = \begin{bmatrix} \frac{\partial L(\beta)}{\partial \beta_0} \\ \dots \\ \frac{\partial L(\beta)}{\partial \beta_{p-1}} \end{bmatrix}$$

Esta é a função *score* para a primeira derivada da função. O vetor de parâmetros inicial que muitos softwares utilizam para calcular o erro das estimativas é zero, ou seja,

$$\beta_{k=0} = [0 \quad \dots \quad 0]$$

Além do que foi explicitado acima, pode-se calcular um intervalo de confiança para cada um dos parâmetros do modelo a um nível de confiança  $(1-\alpha)\%$  através da equação (4.9):

$$\hat{\beta}_j \pm z_{1-\frac{\alpha}{2}} \sigma(\hat{\beta}_j)$$

#### 4.1.4- Testes e Validação do Modelo

Após o cálculo dos coeficientes, o próximo passo consiste em determinar se as variáveis independentes são associadas significativamente à variável resposta. Os métodos inferenciais nos modelos lineares generalizados baseiam-se, principalmente, na teoria da máxima verossimilhança.

A função de verossimilhança pode ser avaliada para diferentes  $\theta$ . No entanto, a estimação dos parâmetros  $\theta$  deve seguir a premissa de maximização da função de verossimilhança, ou seja, a estimação deve partir da maximização dos valores em cada um dos experimentos. Segundo LOUVIERE et al (2000), tal consideração leva à seguinte regra para estimação por máxima verossimilhança: o estimador de máxima verossimilhança de  $\theta$  é o valor  $\bar{\theta}$  que maximiza a função  $L^*(\theta)$ .

Maximizar a função de verossimilhança exige muito esforço, sendo mais conveniente maximizar a função  $\log L^*(\theta)=L$ , como ilustrado em (4.10). Sendo assim,

$$MaxL(\theta) = Max \sum_{t=1}^T P_{x_t} \left( \frac{x_t}{\theta} \right) \quad (4.10)$$

Neste caso, tem-se que:

$$\sum_{t=1}^T \frac{\partial}{\partial \theta_k} \ln P_{x_t} \left( \frac{x_t}{\theta} \right) = 0, \quad k=1,2,3,\dots,r$$

A estimação dos parâmetros  $\theta$  pode ser feita através de um processo iterativo de gradiente de busca, chamado de método de *Newton-Raphson*, demonstrado por LOUVIERE et al (2000) e RUGGIERO e LOPES (1996).

Feita a calibração pelo método de ajuste por máxima verossimilhança, é possível estimar alguns termos estatísticos essenciais na análise de *performance* do modelo estimado. Os parâmetros de *performance* originadas da estimação dos modelos são utilizados não somente para a avaliação dos modelos Multinomial Logit, como também para os outros modelos da família *logit*. Os principais são:

- $L(0)$ : valor da função logarítmica de verossimilhança quando todos os parâmetros são zero;
- $L(c)$ : é o valor da função logarítmica de verossimilhança quando somente a constante específica da alternativa é incluída;
- $L(\beta)$ : valor máximo da função logarítmica de verossimilhança;
- $-2(L(0) - L(\beta))$ : estatística utilizada para testar a hipótese nula de que todos os parâmetros são zero; é assintoticamente distribuída com  $K$  graus de liberdade, em que  $k$  é o número de coeficientes do modelo;

- $-2(L(c) - L(\beta))$ : estatística utilizada para testar a hipótese nula de que todos os parâmetros são zero; é assintoticamente distribuída com  $\chi^2$  com  $k-j+1$  graus de liberdade;
- $\rho_{zero}^2$ : informal “goodness of fit”. Mede a fração do valor de verossimilhança explicado pelo modelo, definido como  $1 - \left( \frac{L(\beta)}{L(0)} \right)$ . Os valores de  $\rho_{zero}^2$  dependem do tipo de modelo a ser construído. Essa medida é mais adequada na comparação de duas especificações desenvolvidas com o mesmo conjunto de dados;
- $\rho_{const}^2$ : reflete o grau de convergência na estimação dos parâmetros do modelo com os atributos dos coeficientes, em relação ao modelo estimado somente com as ASC’s. É dado por  $1 - \left( \frac{L(\beta)}{L(c)} \right)$ ;
- $\overline{\rho}^2$ : valor similar ao  $\rho^2$ , porém corrigido para números de parâmetros estimados.  $\overline{\rho}^2 = 1 - \left( \frac{L(\beta) - K}{L(0)} \right)$ , onde  $k$ , já definido, é o número de graus de liberdade.

Baseando-se nessas estatísticas, é possível a realização de um teste estatístico para avaliar a hipótese nula de que o modelo com todas as variáveis independentes explica melhor o modelo. O teste mais comum é o proposto por SMALL e HSIAO (1982), descrito em BEN-AKIVA e LERMAN (1985). O teste, semelhante ao da razão de máxima verossimilhança, consiste em estimar uma função de verossimilhança com a retirada de alguma das variáveis independentes e com isso será feito a partir da razão das log-verossimilhanças.

Mais uma estatística de teste que pode ser utilizada é o pseudo-coeficiente de determinação (MCFADDEN, 1974), onde se calcula a relação de informação que é perdida, quando alguma variável do modelo é retirada. Para esse cálculo é utilizada a equação (4.10):

$$\rho^2 = 1 - \frac{\xi(\beta)}{\xi(0)} \quad (4.10)$$

Por fim para o modelo ajustado com a redução dos graus de liberdade do modelo pode-se expressar pela equação (4.11):

$$\rho^2 = 1 - \frac{\xi(\beta) - K}{\xi(0)} \quad (4.11)$$

## 4.2 – Simulações

Em computação, simulação consiste em empregar técnicas matemáticas em computadores com o propósito de imitar um processo ou operação do mundo real. Desta forma, para ser realizada uma simulação, é necessário construir um modelo computacional que corresponda à situação real que se deseja simular. Em outras palavras, é o estudo do comportamento do mundo real através de um modelo.

Existem diversas definições para a simulação, dentre elas, uma citada por diversos autores, em função de sua abrangência, é a de Pegden et al (1990), que afirmam que “a simulação é um processo de projetar um modelo computacional de um sistema real e conduzir experimentos com este modelo com o propósito de entender seu comportamento e/ou avaliar estratégias para sua operação”. Desta maneira, pode-se entender a simulação como um processo amplo que engloba não apenas a construção do modelo, mas todo o método experimental que se segue, buscando:

- Descrever o comportamento do sistema;
- Construir teorias e hipóteses considerando as observações efetuadas;
- Usar o modelo para prever o comportamento futuro, isto é, os efeitos produzidos por alterações no sistema ou nos métodos empregados em sua operação.

Essa definição mostra que a simulação é muito mais do que meramente a construção do modelo de simulação ou então muito mais do que meramente a experimentação, ou simulação propriamente dita, desse modelo. A simulação inclui todo “processo de projetar um modelo computacional de um sistema real”, e isso inclui a formulação do problema, a especificação do funcionamento de tal sistema real, uma análise de riscos, a construção do modelo propriamente e toda sua verificação e validação.

Com os avanços na área de informática, modernos equipamentos e novas linguagens de programação e de simulação têm permitido empregar a técnica de simulação nas diversas áreas do conhecimento humano, como por exemplo: projetar e analisar sistemas industriais, avaliar performance de hardware e software em sistemas de computação, determinar



frequência de pedidos de compra para recomposição de estoques, projetar e administrar sistemas de transportes como: portos e aeroportos, etc.

Já faz alguns anos que o recurso a técnicas de simulação a utilização de números aleatórios é utilizado como alternativa aos procedimentos mais convencionalmente utilizados em estatística para testar hipóteses, Waton e Blackstone (1989). O uso da simulação de sistemas tem crescido muito nas últimas décadas, à medida que o poder computacional disponível também aumenta. Sua aplicação atinge hoje praticamente qualquer ramo da atividade humana, e em várias etapas do desenvolvimento, desde a escolha de alternativas a projetos já existentes até a validação de novas idéias. Segundo Freitas Filho (2010), a simulação permite que “estudos sejam realizados sobre sistemas que ainda não existem, levando ao desenvolvimento de projetos eficientes antes que qualquer mudança física tenha sido iniciada”.

A simulação utilizada no presente faz uso de variáveis aleatórias, geradas a partir da função “aleatório()” do Excel 2003.

O algoritmo do Excel 2003 para gerar números aleatórios foi desenvolvido por B.A Wichman e I.D Hill (1987). Foi demonstrado por Rotz et Al (2001) que esse algoritmo passa nos testes de aleatoriedade dos resultados.

O resultado do modelo logístico múltiplo no presente trabalho é a probabilidade para o resultado de um confronto entre duas equipes, ou seja, vitória do time da casa, empate entre as equipes ou derrota do time que detém o mando de campo.

Esse modelo será calculado até uma determinada rodada e, com base nele, serão simulados os possíveis resultados dos jogos não realizados do campeonato.

A soma das probabilidades do resultado de cada jogo será sempre 1, ou seja, supondo, num jogo entre as equipes A e B, que a probabilidade do jogo terminar com vitória do time da casa seja de 0.43, que a probabilidade do jogo terminar empatado seja de 0.37, significa que a probabilidade do resultado ser favorável ao time visitante será de 0.30.

A utilização da simulação no trabalho servirá para identificar o resultado mais provável de ocorrer, com base no resultado do modelo. Será gerado um número aleatório uniformemente distribuído entre 0 e 1 para os jogos não realizados do campeonato. A partir desse número aleatório e da probabilidade do resultado (gerada a partir do modelo), será atribuído um resultado para a partida, da seguinte forma:

Suponha que a probabilidade de um determinado jogo terminar com vitória do time da casa é de 0.43, e que a probabilidade do jogo terminar empatado é de 0.37, isso significa que a

probabilidade do resultado ser favorável ao time visitante será de 0.30 (resultados obtidos a partir do modelo logístico múltiplo). Para esse confronto (ainda não realizado), será gerado um número aleatório entre 0 e 1 para simular o resultado da partida. O resultado da partida será dado a partir da seguinte classificação: se o número aleatório for entre 0 e 0.43, significa que o resultado será vitória para o time da casa, agora se o número aleatório for maior que 0.43 e menor que 0.60 (0.43+0.37) significa que o resultado associado a partida será o empate, e, se o número aleatório for maior que 0.70, o resultado para a partida será a vitória do time visitante.

Sendo mais geral, para cada partida, temos um número  $p < 1$  que é a probabilidade da vitória do mandante, outro número  $q < 1-p$  que dá a probabilidade do empate. Gera-se um número aleatório  $x$  entre 0 e 1. Se  $x \leq p$ , significa que o mandante vence este jogo nesta simulação; se  $p < x \leq p+q$ , significa que há empate; se  $x > p+q$ , o visitante é que venceu. Entendido como é feita a simulação para uma partida, basta aplicar esse mesmo princípio para todas as partidas não realizadas a partir da rodada que o modelo é estimado. Os números  $p$  e  $q$  que definem as probabilidades para cada jogo são obtidos a partir do modelo logístico múltiplo.

A simulação trabalhada no presente trabalho será calculada com base nas probabilidades estimadas para os jogos não realizados do campeonato, derivadas dos modelos estimados com os jogos realizados até a rodada, isto é, a cada rodada será estimado um modelo e calculada a probabilidade de cada resultado possível para todos os jogos não realizados do campeonato. Em outras palavras, a partir da finalização de uma rodada (quando os jogos são realizados), será estimado um novo modelo (agora considerando também essa rodada) assim como novas simulações serão feitas para os resultados futuros.

A cada jogo não realizado, será atribuído um número aleatório entre 0 e 1 e, com base na probabilidade derivada do modelo e desse número aleatório, será atribuído o resultado ao jogo. Com base nos resultados estimados e nos resultados dos jogos já realizados, serão calculados os pontos estimados no final do campeonato e, através dessa pontuação estimada, analisados os times que figuram entre os resultados de interesse (times rebaixados e classificados para a Copa Libertadores da América) considerando todas as simulações realizadas.

Foram feitos diversos testes com números de simulações distintas com o intuito de avaliar a quantidade de simulações necessárias que não gerasse variações nos resultados. Para alguns modelos, foram feitas 1000, 2000, 3000, 4000, até 10000 simulações (variações de

1000 em 1000) e os resultados finais (classificação para a copa libertadores da América e rebaixamento para a série B) praticamente não sofriram variações. Em função disso, optou-se por trabalhar com 1000 simulações no presente trabalho.

## **5. MODELOS DO PRESENTE TRABALHO**

Agora, serão apresentados os modelos desenvolvidos e toda a parte que envolve as variáveis a serem utilizadas nos mesmos. Em relação às variáveis, é importante ressaltar o uso do *home advantage*, que visa mostrar a vantagem que o time tem ao ter os jogos realizados em sua casa, pois na maioria dos esportes o mando de campo é algo crucial para o bom desempenho de um time no campeonato.

Além do *home advantage*, serão tratadas variáveis como desempenho recente do time, pois este ponto é importante em diversos aspectos para melhorar o rendimento de uma equipe. Logo a partir de agora serão explicados e desenvolvidos todos os modelos que irão ser trabalhados no presente estudo.

O presente trabalho não utiliza as probabilidades provenientes do modelo como informações preditivas. A idéia é combinar o resultado do modelo de previsão com técnicas de simulação para avaliar a possibilidade das equipes se classificarem para a Copa Libertadores da América e serem rebaixadas.

Todos os modelos utilizam a força da equipe como variável dependente, essa força da equipe é trabalhada da seguinte maneira: se a equipe está jogando em casa, a variável X (dependente) para essa equipe receberá “+1” e se a equipe está jogando como visitante, essa variável recebe “-1”. Desta forma, os coeficientes “betas” estimados para essas variáveis funcionam como uma força individual de cada time, enquanto o intercept agrega ao time da casa uma força além do que a equipe possui. Esse intercept será interpretado como a vantagem de jogar em casa.

## 5.1 – Descrição dos Modelos

Dentre os modelos testados para a previsão dos resultados do campeonato brasileiro de 2008, três mereceram destaque:

- a. Modelo 1 - força do time e vantagem de jogar em casa → É o modelo mais simples dentre os testados no presente trabalho. Esse modelo trabalha com duas informações para gerar a probabilidade do resultado dos jogos não realizados no campeonato: a força individual de cada time e a vantagem do time em jogar em casa. Uma vantagem desse modelo é a simplicidade, entretanto, esse modelo dá a mesma importância a todos os jogos do campeonato, o que não é verdade por diversos fatores, dentre os quais vale destacar: foco de mais de uma competição, aquisição de novos jogadores, mudança de técnico, etc. Em função da inércia observada nesse modelo, ou seja, na demora em captar mudanças de comportamento das equipes ao longo do campeonato, foram testados os modelos 2 e 3.
- b. Modelos 2 - força do time, vantagem de jogar em casa e desempenho recente dentro e fora de casa → Além de considerar a força individual de cada time e a vantagem de jogar em casa, foi incorporada uma nova informação, com o objetivo de captar o desempenho recente da equipe, que está sendo trabalhado considerando os três últimos jogos realizados em casa (ou fora de casa). Essa informação é importante em função de alguns times apresentarem melhores desempenhos bastante distintos como mandantes ou como visitantes. Como exemplo desse feito, pode-se destacar a equipe do Sport do Recife, que venceu a copa do Brasil de 2008 basicamente com jogos realizados na ilha do retiro (que, diante dessa campanha em casa, foi “batizada” de “la bombonilha”, em alusão ao estádio da equipe do argentina do Boca Juniors, “La Bombonera”, onde, essa equipe dificilmente sofre derrotas.
- c. Modelo 3 - força do time, vantagem de jogar em casa e desempenho recente do time → Além de considerar a força individual de cada time e a vantagem de jogar em casa, foi incorporada uma nova informação, com o objetivo de captar o desempenho recente do time. Entretanto, esse desempenho não terá relação com o

time ter jogado em casa ou fora de casa, ou seja, a diferença entre esses modelos e o descrito anteriormente se dá pela maneira que o desempenho recente é incorporado ao modelo; nesses modelos essa diferença independe de onde o time jogou. Em ambos os modelos 2 e 3, não está sendo utilizada uma série histórica, o que está sendo feito é incorporar uma variável que capta uma mudança recente de desempenho da equipe.

### 5.1.1 – Modelo 1 – Força do Time e Vantagem de Jogar em Casa

Esse modelo tenta prever a probabilidade do resultado dos jogos não realizados a partir dos jogos realizados. Como informações preditivas foram utilizadas a vantagem de jogar em casa e a força individual do time.

A vantagem de jogar em casa vem sendo tratada por diversos autores em diversos esportes e o impacto dessa característica no resultado final dos jogos é defendida por diversos autores, sendo aplicada tanto no futebol como no beisebol, voleibol, futebol americano, entre outros esportes (Courneya e Carron, 1992; Pollard, 2002).

Segundo Courneya e Carron (1992) e Pollard (2002), como fatores que contribuem para a vantagem de jogar em casa, destaca-se a familiaridade com o campo e estádio, isso porque há uma maior consciência do atleta da "casa", o que lhe permite uma orientação mais eficaz nas ações decorrentes e exigidas na partida.

Lawal e Sundheim (2007) utilizaram essas mesmas informações para prever o resultado do EPL (English Premier League) de 2002; ou seja, utilizaram a vantagem de jogar em casa e as forças individuais de cada time como variáveis preditivas.

O modelo proposto é:

$$\ln \left( \frac{P_{ij}}{1 - P_{ij}} \right) = \alpha_i + \sum_{t=1}^{20} \phi_t X_t$$

Onde:

$P_{ij}$  é a probabilidade do time "i" vencer o time "j" jogando em casa

$\alpha_i$  é o *Home Advantage*

$$X_t = \begin{cases} 1 & (\text{time está em casa}) \\ -1 & (\text{time é visitante}) \\ 0 & (\text{se não está jogando}) \end{cases}$$

Esse modelo possui uma restrição em sua variável resposta, que precisa ser binária.

Com o objetivo de utilizar melhor a informação do resultado dos jogos, optou-se por trabalhar com um modelo *logito* Ordinal, da seguinte forma – já adaptado para 20 times:

$$\ln \left[ \frac{P(\text{Vitória}/X_1 - X_{19}, \delta)}{P(\text{Empate ou Derrota} / X_1 - X_{19}, \delta)} \right] = \alpha_1 + \sum_{i=1}^{20} \phi_i X_i$$

$$\ln \left[ \frac{P(\text{Vitória ou Empate} / X_1 - X_{19}, \delta)}{P(\text{Derrota} / X_1 - X_{19}, \delta)} \right] = \alpha_2 + \sum_{i=1}^{20} \phi'_i X_i$$

Esse modelo está sujeito a:

$$\phi'_i = \phi_i \quad \forall i = 1, 2, \dots, 20$$

Ou seja, a força do time “i” jogando contra o time “j” é a mesma, independente se a variável resposta é apenas a vitória ou uma vitória ou empate em casa.

Desta forma, o modelo final fica:

$$\ln \left[ \frac{P(\text{Vitória}/X_1 - X_{19}, \delta)}{P(\text{Empate ou Derrota} / X_1 - X_{19}, \delta)} \right] = \alpha_1 + \sum_{i=1}^{20} \phi'_i X_i$$

$$\ln \left[ \frac{P(\text{Vitória ou Empate} / X_1 - X_{19}, \delta)}{P(\text{Derrota} / X_1 - X_{19}, \delta)} \right] = \alpha_2 + \sum_{i=1}^{20} \phi'_i X_i$$

Onde:

$\alpha_1$  é a vantagem de jogar em casa quando o interesse é a vitória

$\alpha_2$  é a vantagem de jogar em casa quando o interesse é a vitória ou o empate

$\phi'_i$  é a força individual de cada time

O modelo utiliza uma variável binária para identificar a partida, da seguinte maneira, se a equipe está jogando em casa, a variável dependente recebe “+1” e se a equipe está jogando como visitante essa variável dependente recebe “-1”. Desta forma, os coeficientes “betas” estimados para as equipes funcionam como uma força individual de cada time, enquanto o intercept agrega ao time da casa uma força além do que a equipe possui. Esse intercept será interpretado como a vantagem de jogar em casa. Vale destacar que, em nenhum modelo testado no presente trabalho, o intercept apresentou coeficientes negativos.

Através desse modelo, é possível calcular as probabilidades de vitória, empate ou derrota para um time jogando em casa, isso é feito com base na fórmula:

$$P(R \leq j) = \left[ \exp(\alpha'_j + \sum_{i=1}^{20} \phi'_i X_i) \right] / \left[ \exp(1 + \alpha'_j + \sum_{i=1}^{20} \phi'_i X_i) \right] \quad (5.1)$$

Com base nesse modelo, pode-se isolar o intercept do modelo, ou seja, estimar a força do time em um confronto direto contra os demais times, criando, desta forma, um *ranking* dos times, isso é feito com base na fórmula abaixo:

$$P_i = [\exp(\theta'_i)] / \left[ \sum_{i=1}^{20} \exp(\theta'_i) \right] \quad (5.2)$$

Um dos problemas desse modelo é a importância dada a cada jogo, ou seja, todos os jogos são igualmente importantes para definir a probabilidade dos jogos futuros. Em função disso, esse modelo demora muito para captar as mudanças de comportamento de uma equipe. Fato este que fez com que fossem trabalhados mais dois modelos, com o intuito de captar mais rapidamente uma mudança de desempenho das equipes.

Para estimar os coeficientes do modelo, foi utilizado o software SPSS 16.

### 5.1.2 – Modelo 2 - Força do Time, Vantagem de Jogar em Casa e Desempenho Recente Dentro e Fora de Casa

O objetivo desse modelo é captar mais rapidamente uma mudança de desempenho da equipe, pois o modelo anterior levava muito tempo para captar essa mudança. Desta forma, esse modelo acrescenta uma informação ao modelo anterior, que é o desempenho recente do time, porém, esse desempenho foi pensado de maneira a tratar a realidade do time ao jogar em casa e a força dele quando era time visitante, ou seja, essa variável traz o desempenho recente do time dentro de casa ou fora de casa. Faz-se necessário esse teste, pois algumas equipes apresentam um desempenho quando jogam fora de seus domínios muito aquém do que conseguem atingir quando jogam em casa.

Como equipes que apresentam um desempenho muito superior em casa, vale destacar a equipe do Sport, que se sagrou campeã da copa do Brasil de 2008 a partir de vitórias em casa.

A variável de desempenho recente do time ( $Z_i$ ) é calculada com base nos três últimos jogos da equipe, da seguinte maneira. Do total de pontos que a equipe poderia alcançar nos três últimos jogos (total de 9 pontos), qual foi o percentual que a equipe alcançou?

Por exemplo, haverá um confronto entre as equipes A e B (A jogando em casa), a variável  $Z_i$  para esse confronto era calculada da seguinte maneira:

Divide-se o total de pontos que a equipe A alcançou nos três últimos jogos que realizou sendo mandante do campo (jogos em casa) por 9 (que é o total de pontos possível). Faz-se o mesmo para a equipe visitante, ou seja, calcula-se o percentual de pontos que essa equipe alcançou nos três últimos jogos em que foi visitante.



A variável  $Z_i$  é a diferença entre a força do time que está jogando em casa pelo time visitante. Desta forma, o time A recebe  $+Z_i$  e o time B recebe  $-Z_i$ , ou seja, o desempenho recente das equipes (jogando em casa e fora de casa) está sendo considerado para ambos os times.

Desta forma, esse modelo considera, em média, seis rodadas já realizadas pela equipe, isso porque a tabela do campeonato é definida de maneira a haver uma alternância entre os jogos, ou seja, joga-se um jogo em casa e o seguinte fora de casa. Em função das equipes participarem de mais de um campeonato ao mesmo tempo, eventualmente esse cronograma pode sofrer alterações.

Ao acrescentar essa variável, o modelo final fica:

$$\ln \left[ \frac{P(\text{Vitória} / X_1 - X_{19}, \delta)}{P(\text{Empate ou Derrota} / X_1 - X_{19}, \delta)} \right] = \alpha_1 + \sum_{i=1}^{20} \phi_i X_i + \sum_{i=1}^{20} \beta_i Z_i$$

$$\ln \left[ \frac{P(\text{Vitória ou Empate} / X_1 - X_{19}, \delta)}{P(\text{Derrota} / X_1 - X_{19}, \delta)} \right] = \alpha_2 + \sum_{i=1}^{20} \phi'_i X_i + \sum_{i=1}^{20} \beta'_i Z_i$$

Os coeficientes estimados para o modelo significam:

$\alpha_1$  é a vantagem de jogar em casa quando o interesse é a vitória

$\alpha_2$  é a vantagem de jogar em casa quando o interesse é a vitória ou o empate

$\phi_i$  é a força individual de cada time

$\beta_i$  é a força do desempenho recente jogando em casa ou fora de casa

Assim como no modelo anterior, é possível calcular a força individual das equipes, com base na equação 5.2.

### 5.1.3 – Modelo 3 - Força do Time, Vantagem de Jogar em Casa e Desempenho Recente do Time

Outra maneira de captar a mudança recente do desempenho do time seria desvincular esse desempenho recente do local do jogo, ou seja, independente do time ter jogado em casa ou fora de casa, ele apresentou uma melhora nos últimos jogos ou não.

Assim como no modelo anterior, o interesse do modelo três é acrescentar uma informação capaz de captar o desempenho recente da equipe ao modelo. Entretanto, nesse modelo, essa informação será calculada com base no desempenho da equipe nas últimas 6 partidas, independente da equipe ter jogado em casa ou fora de casa.

A quantidade de jogos foi mantida em 6 com o intuito de fazer com que esse modelo mantenha, em média, um mesmo número de jogos que o modelo 2, tornando, assim, as simulações resultantes de ambos os modelos mais próximas, por considerarem um mesmo número de rodadas na formação do desempenho recente.

Esse modelo é interessante e precisa ser testado, pois, as equipes podem apresentar um desempenho recente fantástico independente de onde esteja jogando; como, por exemplo, a equipe do Flamengo no campeonato Brasileiro de 2008, que conseguiu atingir 20 pontos, dos 24 disputados, nas oito primeiras rodadas do campeonato.

Para simplificar a maneira que essa variável é calculada, toma-se o seguinte exemplo: suponha que haverá um confronto entre as equipes A e B (A jogando em casa), a variável  $Q_i$  para esse confronto era calculada da seguinte maneira:

Divide-se o total de pontos que a equipe A alcançou nos seis últimos jogos que realizou sendo mandante do campo (jogos em casa) por 18 (que é o total de pontos possível, ou seja, vitória em todos os 6 jogos). Faz-se o mesmo para a equipe visitante, ou seja, calcula-se o percentual de pontos que essa equipe alcançou nos seis últimos jogos.

A variável  $Q_i$  é a diferença entre a força do time que está jogando em casa pelo time visitante. Desta forma, o time A recebe  $+Q_i$  e o time B recebe  $-Q_i$ , ou seja, o desempenho recente das equipes (independente de onde foi realizado o jogo) está sendo considerado para ambos os times. Faz-se necessário manter um número par de jogos para que a vantagem de jogar em casa não interfira.

Ao acrescentar essa variável, o modelo final fica:

$$\ln \left[ \frac{P(\text{Vitória} / X_1 - X_{19}, \delta)}{P(\text{Empate ou Derrota} / X_1 - X_{19}, \delta)} \right] = \alpha_1 + \sum_{i=1}^{20} \phi_i X_i + \sum_{i=1}^{20} \varphi_i Q_i$$

$$\ln \left[ \frac{P(\text{Vitória ou Empate} / X_1 - X_{19}, \delta)}{P(\text{Derrota} / X_1 - X_{19}, \delta)} \right] = \alpha_2 + \sum_{i=1}^{20} \phi'_i X_i + \sum_{i=1}^{20} \varphi'_i Q_i$$

Os coeficientes estimados para o modelo significam:

$\alpha_1$  é a vantagem de jogar em casa quando o interesse é a vitória

$\alpha_2$  é a vantagem de jogar em casa quando o interesse é a vitória ou o empate

$\phi_i$  é a força individual de cada time

$\varphi_i$  é a força do desempenho recente da equipe

Assim como no modelo anterior, é possível calcular a força individual das equipes, com base na equação 5.2.

## **6. ANÁLISE DE RESULTADOS**

Neste capítulo serão feitas as análises dos resultados obtidos com os modelos, porém o capítulo começará com algumas análises básicas dos resultados do campeonato, que ajudarão a elucidar o uso de algumas variáveis nos modelos.

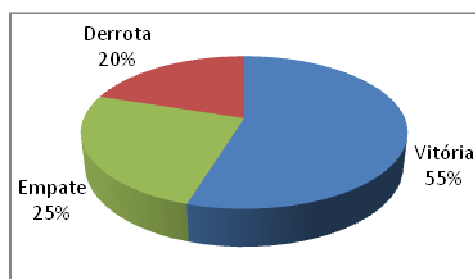
Na segunda parte serão feitas análises relativas aos 3 modelos desenvolvidos, com destaque para as nuances entre eles.

### **6.1 – Análise Exploratória dos jogos**

Essas análises serão feitas com base no campeonato brasileiro como um todo, com destaque para o desempenho dos times em casa e fora de casa, distribuição dos resultados das partidas, assim como o desempenho individual de cada equipe tanto em casa quanto fora de casa.

Na figura 6.1 tem-se a distribuição de todos os jogos do campeonato conforme os 3 tipos de resultados possíveis: vitória do mandante, empate e vitória do visitante.

Figura 6.1 – Distribuição dos resultados das partidas pelo time que está jogando em casa

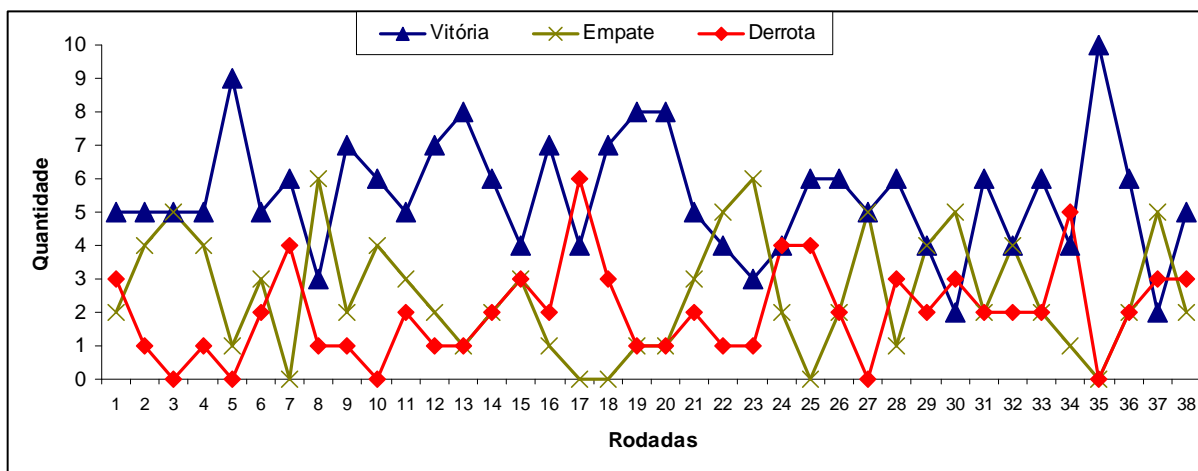


Pode-se perceber que no campeonato de 2008 em quase 55% dos jogos o time da casa fez prevalecer o mando de campo, o que mostra que no futebol brasileiro ter um bom desempenho em casa é fundamental para um bom desempenho no final do campeonato.

Alguns modelos utilizam o empate como forma de considerar que o time tem força em casa, pois, por este prisma, pode-se dizer que o time se mantém invicto. Ao observar a figura 6.1, pode-se concluir que neste campeonato os times visitantes tiveram muitas dificuldades, pois em apenas 20% das partidas conseguiram suplantar o adversário.

Na figura 6.2 estão apresentados os desempenhos rodada a rodada em relação à pontuação obtida pelos times da casa e pelo time de fora, ou seja, em cada rodada temos 10 jogos, quantos desses jogos tiveram vitória do time da casa, quantos empataram e quantos o time visitante venceu.

Figura 6.2 – Desempenho dos mandantes e visitantes



Na figura 6.2 observa-se que em 33 rodadas os times da casa conseguem obter mais vitórias que os times visitantes, isto mostra que existe uma grande importância o mando de campo e isto pode ser visto também na tabela 6.1, que mostra a pontuação total de cada equipe, assim como a pontuação adquirida em jogos como mandante e como visitante.

Considerando as 38 rodadas do campeonato, em somente 3 delas houve mais vitórias dos times visitantes do que para os times mandantes, isso ocorreu nas rodadas 17, 30, 34 e 37. O extremo ocorreu na rodada 35, onde todos os times que possuíam o mando de campo se sagraram vitoriosos nas 10 partidas disputadas.

A tabela 6.1 reforça a importância do time sair vitorioso no mando de campo para se manter competitivo no campeonato.

Tabela 6.1: Pontuação total e % dos pontos obtidos como mandante e como visitante

	Total Pontos 2008	% Pontos obtido Em Casa	% Pontos obtido Fora de Casa
São Paulo	75	61,3%	38,7%
Grêmio	72	63,9%	36,1%
Cruzeiro	67	70,1%	29,9%
Palmeiras	65	67,7%	32,3%
Flamengo	64	57,8%	42,2%
Internacional	54	79,6%	20,4%
Botafogo	53	58,5%	41,5%
Coritiba	53	67,9%	32,1%
Goiás	53	69,8%	30,2%
Sport	52	75,0%	25,0%
Vitória	52	73,1%	26,9%
Atlético_MG	48	72,9%	27,1%
Atlético_PR	45	80,0%	20,0%
Fluminense	45	60,0%	40,0%
Santos	45	73,3%	26,7%
Figueirense	44	59,1%	40,9%
Náutico	44	72,7%	27,3%
Vasco	40	62,5%	37,5%
Portuguesa	38	84,2%	15,8%
Ipatinga	35	85,7%	14,3%

A tabela 6.1 reforça que, para um bom desempenho final, faz-se necessário, além de uma grande campanha em casa, uma boa campanha fora de casa, ou seja, a importância de cada jogo nesse formato de campeonato. Pode-se tomar como exemplo a equipe do Internacional, que teve uma campanha excelente em casa, porém conseguiu apenas 11 pontos (20.4% dos seus 54 pontos) como visitante, o que atrapalhou o time na disputa por objetivos maiores.

Um time que sofreu com seu desempenho como mandante foi o Flamengo que fez 10 pontos a menos do que o Cruzeiro, que teve a melhor campanha como mandante. Fato este que contribuiu para que o Flamengo não conseguisse a vaga na Copa Libertadores da América, pois ele só ficou a 1 ponto da classificação.

Na parte de baixo da tabela, ou seja, dos times que foram rebaixados, observa-se que as equipes têm muita dificuldade de conseguir pontuar fora de casa, pois os 2 últimos

classificados (Ipatinga e Portuguesa) conseguiram apenas 5 e 6 pontos respectivamente, o que influenciou de forma bastante grande no rebaixamento destas equipes.

Apesar de apresentar um bom desempenho fora de casa, o Vasco foi um dos times com o menor aproveitamento dos jogos em casa, assim como o Figueirense que também foi rebaixado para a série B em 2009.

Um último ponto que tem importância e precisa ser destacado são os jogos em casa, onde não existe um time que possa ser considerado mandante, estes jogos existem principalmente em estados onde existe um estádio de grande porte e com isso os chamados clássicos regionais são disputados em um campo que pode ser considerado neutro. As partidas que poderiam ser consideradas em campo neutro são:

- Flamengo x Fluminense;
- Flamengo x Vasco;
- Flamengo x Botafogo;
- Fluminense x Vasco;
- Fluminense x Botafogo;
- Vasco x Botafogo;
- Cruzeiro x Atlético-MG.

Atualmente, os jogos realizados em São Paulo não estão entre os exemplos, pois o Palmeiras voltou a realizar os clássicos regionais no Palestra Itália, vulgo Parque Antarctica, e o São Paulo como mandante organiza os jogos no Morumbi.

## 6.2 – Análise dos modelos

Neste capítulo serão feitas as análises relativas aos modelos desenvolvidos e explicados previamente. Este capítulo será sub-dividido em 4 menores, onde os 3 primeiros contemplarão as análises referentes a cada um dos modelos individualmente e, uma última parte com algumas comparações entre os resultados obtidos pelos 3 modelos.

Vale ressaltar que diante das inúmeras análises que poderiam ser feitas sobre as previsões e os resultados, o presente trabalho focará na força dos times, na chance de um time conseguir a vaga para a competição Libertadores da América e por fim a chance de rebaixamento. Outro ponto é que as análises serão feitas para algumas equipes, pois uma análise individual para cada time demandaria muito tempo e o texto poderia se tornar bastante cansativo e extenso.

### 6.2.1 – Modelo 1 - Força do Time, Vantagem de Jogar em Casa

Após a breve explicação que todos os modelos tiveram no capítulo anterior, seguem abaixo as análises dos resultados obtidos com a aplicação do modelo nos jogos do campeonato brasileiro de futebol de 2008.

Com o objetivo de ilustrar passo a passo as contas realizadas para se obter a força das equipes, assim como a previsão de um confronto direto entre duas equipes, selecionou-se os coeficientes da rodada 20 desse modelo para cálculo de todas as informações. Seguem abaixo os coeficientes:

Tabela 6.2.1: Coeficientes do Modelo 1 para a rodada 20

	Rodada 20
Intercept Vitória	0,50
Intercept Vitória+Empate	1,86
Atlético_MG	-0,18
Atlético_PR	-0,73
Botafogo	0,70
Coritiba	0,59
Cruzeiro	0,95
Figueirense	0,10
Flamengo	0,48
Fluminense	-0,93
Goiás	-0,25
Grêmio	1,42
Internacional	-0,03
Ipatinga	-1,13
Náutico	-0,83
Palmeiras	0,92
Portuguesa	-0,71
Santos	-0,74
São Paulo	0,59
Sport	-0,07
Vasco	-0,59
Vitória	0,42

Com base nos coeficientes dessa rodada, pode-se calcular a vantagem do time da casa vencer a partida com base nos jogos realizados até a rodada 20, independente de quais foram os confrontos realizados, da seguinte maneira:

$$Home\_Advantage_{vitória} = \frac{e^{\alpha_{vitória}}}{1 + e^{\alpha_{vitória}}}$$

Substituindo pelo intercept da vitória do time da casa, tem-se:

$$Home\_Advantage_{vitória} = \frac{e^{0.5}}{1 + e^{0.5}} = 0.60$$

Ou seja, considerando os jogos realizados até a rodada 20, pode-se afirmar que o time da casa tinha uma vantagem de sair vitorioso do confronto em 60% dos casos.

Analogamente, pode-se analisar a possibilidade do time da casa conseguir pelo menos um ponto, ou seja, vencer ou empatar no confronto, com base no intercept da vitória ou empate, como segue abaixo:

$$Home\_Advantage_{vitória\_empate} = \frac{e^{1.86}}{1 + e^{1.86}} = 0.86$$

Desta forma pode concluir que, com base nos resultados dos jogos realizados até a rodada 20, em 86% dos casos o time da casa conseguia atingir pelo menos um ponto. Ao comparar os resultados obtidos com base no modelo com os resultados reais, pode-se perceber que o modelo está muito bem ajustado à realidade, como mostra a tabela abaixo.

Tabela 6.2.1.2: Resultados para o time com mando de campo até a rodada 20

	%
Vitória	0.600
Empate	0.225
Derrota	0.175

Vale destacar que o modelo encontrou o mesmo valor de vitória para o time da casa que ocorreu no campeonato, assim como um valor bastante próximo para o empate, o que ilustra um bom ajuste do modelo à realidade.

Pode-se calcular a probabilidade de vitória, empate ou derrota de um confronto direto entre duas equipes, da seguinte maneira. Suponha que haverá um confronto entre o Flamengo e o São Paulo, com o Flamengo detendo o mando de campo. A probabilidade de cada um dos resultados possíveis, com base no modelo estimado com base nos jogos realizados até a rodada 20 é de:

$$Vitória(rodada\ 20)_{Flamengo\ x\ São\ Paulo} = \frac{e^{(0.5+0.49-0.59)}}{1 + e^{(0.5+0.49-0.59)}} = 0.59$$

Ou seja, 59% é a probabilidade do Flamengo (com mando de campo) vencer o São Paulo, considerando os jogos realizados até a rodada 20.

Analogamente, pode-se calcular a probabilidade desse jogo ser empate, da seguinte forma:

$$Empate(rodada\ 20)_{Flamengo\ x\ São\ Paulo} = \frac{e^{(1.86+0.49-0.59)}}{1 + e^{(1.86+0.49-0.59)}} = 0.85$$

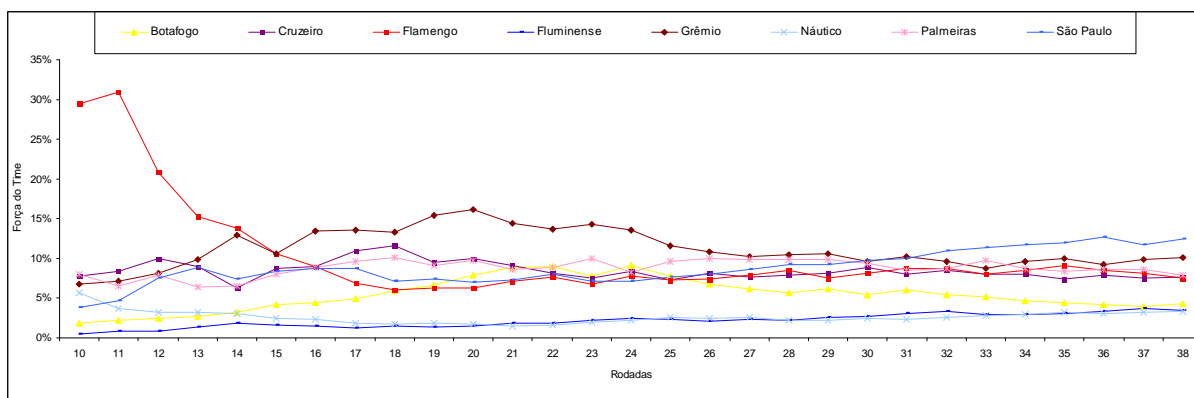


Por fim, a probabilidade da equipe do São Paulo vencer será de 0.15. O cálculo para essas probabilidades é feito de maneira análoga nos demais modelos. A partir de agora serão trabalhados somente os resultados, sem cálculos individuais.

Vale ressaltar que o modelo fornece alguns resultados iguais a 100%, o que não significa certeza absoluta. Os dados são provenientes de 1.000 simulações, desta forma, 100% deve ser interpretado como probabilidade menor do que 1/1000, ao invés de certeza.

O gráfico 6.3 ilustra a variação da força das equipes ao longo das rodadas, ou seja, em cada uma das rodadas é calculado os coeficientes de cada equipe que faz parte do modelo (cálculo feito a partir da equação 5.2) e em seguida é feito o cálculo da probabilidade de vitória em função do coeficiente obtido na equação referente a rodada estudada. Esse cálculo poderia ser interpretado como um *share* das equipes, ou seja, a participação percentual da força das equipes em relação à força total (100%). Em outras palavras, é a força da equipe, independente de qualquer outro fator (vantagem de jogar em casa, desempenho recente, etc). Por exemplo, na rodada 10, a equipe do Flamengo era a que detinha a maior força individual. Isso porque tinha 8 vitórias em 10 jogos realizados.

Figura 6.3 – Força dos Times - Modelo 1



Pode-se perceber através da figura 6.3 que existe uma evolução dos times durante a competição e isto é uma informação bastante importante, pois como todo campeonato longo, existe uma grande oscilação no desempenho das equipes.

Vale ressaltar algumas equipes que tiveram um bom início de campeonato, porém não conseguiram se manter fortes durante toda a competição, um exemplo de grande oscilação pode ser o Flamengo e o Náutico que tiveram um início de campeonato muito bom, porém, após um determinado número de rodadas jogadas, começam a obter resultados ruins, perdendo força no campeonato.

Por outro lado, o São Paulo apresentou um desempenho muito ruim no início de campeonato, principalmente por estar disputando a Copa Libertadores da América. Por este motivo, deixava o campeonato brasileiro em segundo plano, porém, a partir do momento em que foi eliminado da copa Libertadores da América (nas quartas de final, pela equipe do Fluminense); foi aumentando sua força até o momento em que ficou sendo o time com a maior força, que ocorreu a partir da trigésima segunda rodada.

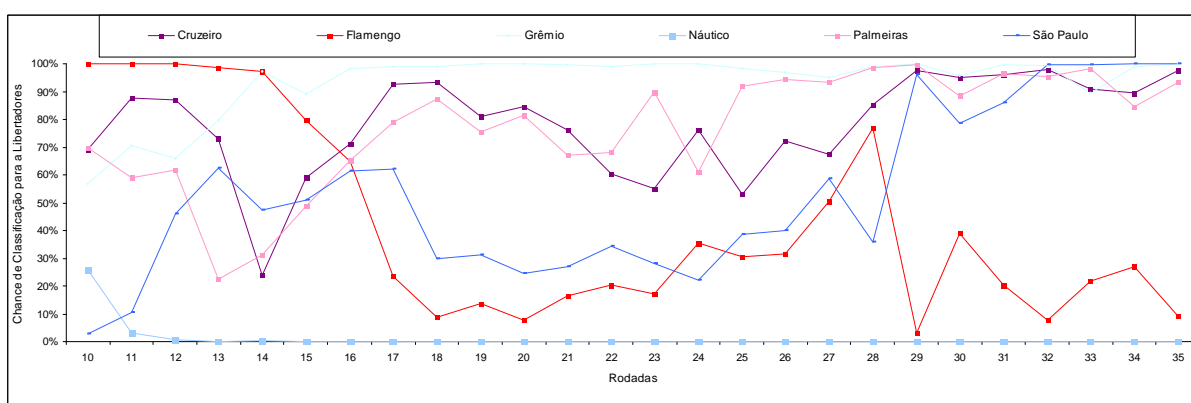
O Fluminense também é um caso que o time teve um aumento na sua força, porém apenas o suficiente para se livrar do rebaixamento, pois ele demorou muito para começar a se dedicar ao campeonato, em função de ter chegado à final da copa Libertadores da América. O Fluminense foi derrotado na final da competição pelo LDU, que é uma equipe equatoriana.

Vale destacar também alguns times que mantiveram um rendimento bastante regular durante toda a competição, conseguindo assim, encerrar a competição entre os 4 times de maior pontuação, são as equipes: Grêmio; Palmeiras e Cruzeiro.

Serão analisadas agora as chances de cada time conseguir a vaga na Libertadores, nesta parte só serão calculadas as chances até a trigésima quinta rodada, pois a partir desta rodada existe muito pouca oscilação, em função da quantidade de jogos a serem realizados.

Essa tabela é calculada a partir de 1000 simulações dos resultados futuros, com base nos jogos realizados até a rodada em questão. As chances de classificação para Libertadores estão expostas na Figura 6.4:

Figura 6.4 – Chances de Classificação para a Copa Libertadores da América rodada a rodada - Modelo 1



Ao observar a figura 6.4, pode-se perceber que mesmo contendo 20 equipes no campeonato, a disputa por uma vaga na Libertadores acaba se resumindo a apenas 5 equipes a partir da trigésima rodada, onde a quinta equipe possuiu uma chance menor do que as demais concorrentes.

Para a vaga na copa Libertadores da América, pode-se perceber que, a partir da décima quinta rodada, quem seriam os candidatos a estas vagas, pois apenas 8 equipes tinham chances razoáveis de alcançar uma das vagas e dentre essas chamam mais atenção 3 que são: Cruzeiro, Palmeiras e Grêmio. Estas equipes mantiveram uma grande regularidade durante toda a competição e sempre figuraram dentro ou muito próximo ao G-4, nome dado aos postulantes a vaga.

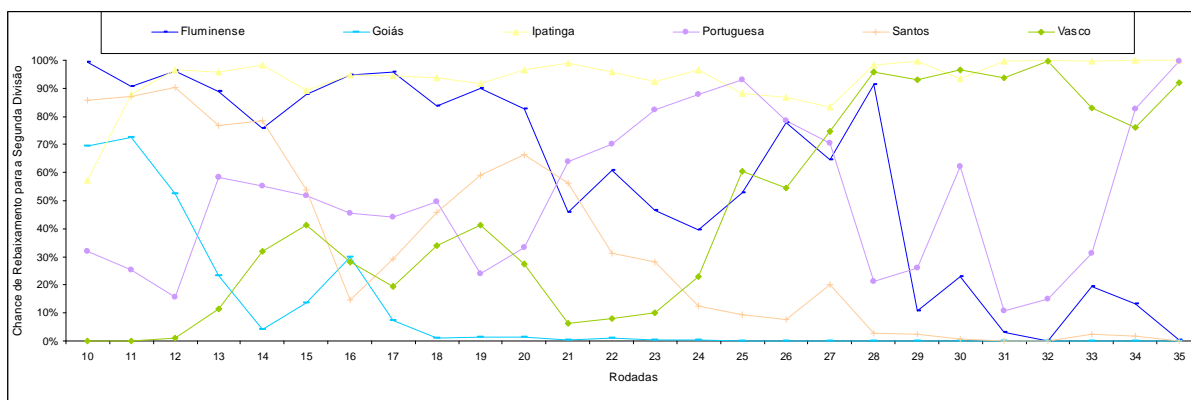
Um time que teve uma grande queda foi o Flamengo, pois até a décima quarta rodada tinha pelo menos 90% de chance em figurar em os 4 de maior pontuação ao final da competição, porém a partir desta rodada o modelo detecta a queda de rendimento do time, que fez um início de campeonato excepcional, onde conseguiu 20 pontos em 24 disputados. A equipe do Flamengo perdeu alguns de seus principais jogadores para outras equipes como Marcinho, Souza e Renato Augusto e não conseguiu repor estes jogadores, acabou caindo de rendimento.

Outro time que teve uma queda de rendimento considerável foi o Náutico que no início da competição fez bons jogos e era um dos postulantes a figurar entre os 4 de maior pontuação, porém, após uma sequência de derrotas e maus resultados, passou a ser um das equipes que brigavam para não ser rebaixados ao final do campeonato.

Por fim vale destacar o São Paulo que tinha chances remotas de conseguir a vaga (e mais remota ainda de ser campeão), mas, com uma ótima sequência de resultados no segundo turno do campeonato (sendo derrotado somente no primeiro jogo – rodada 20), pois a partir desta partida ficou invicto até o final do campeonato e se sagrou campeão da competição.

A última análise será em relação às equipes que irão fazer parte do Z-4, ou seja, as quatro últimas colocações do campeonato, que são as posições menos desejadas por todas as equipes, pois estas posições rebaixam a equipe para a segunda divisão do campeonato, o que impacta de forma substancial nas receitas das equipes. Tudo na série A do campeonato é muito maior, desde os valores dos patrocínios até as cotas de televisão. Para isso, a figura 6.5 ilustra as chances de rebaixamento de cada equipe para a segunda divisão do campeonato brasileiro.

Figura 6.5 – Chances de Rebaixamento para a Segunda Divisão - Modelo 1



Na figura 6.5, pode-se perceber que três equipes iniciam o campeonato de forma bastante ruim e por isso estão com grandes chances de rebaixamento, estas equipes são: Fluminense, Goiás e Santos.

A situação da equipe do Fluminense era bastante clara, porque tinha todos os seus esforços voltados para a disputa da copa Libertadores da América e, por isso, utilizava seu time reserva e não conseguia ter boa campanha.

Já as duas outras equipes caminhavam de forma bastante ruim, pois os times não conseguiam render dentro de campo. O motivo da melhora do Santos foi um crescimento e um bom desempenho em casa principalmente no segundo turno e também ao faro de gol de Kleber Pereira que foi um dos artilheiros do campeonato.

A equipe do Goiás teve a melhora no desempenho em função da contratação do técnico Helio dos Anjos e do jogador Iarley, que veio do Internacional. Estas duas contratações mudaram a cara do time, melhorando assim o seu rendimento.

Dois equipes tiveram um flerte bastante longo com o rebaixamento, que foram Portuguesa e o Ipatinga. Estas equipes permaneceram durante quase todo o campeonato no Z-4, ou seja, sempre tiveram muita chance de rebaixamento.

O Vasco é um time que merece destaque negativo, pois começou o campeonato de forma razoável, conseguindo vitórias e resultados positivos, porém, passou um momento difícil durante o meio do campeonato, o que culminou com a saída do então Presidente Eurico Miranda do cargo e conseqüente rebaixamento da equipe ao final do campeonato. Este rebaixamento pode ser explicado pelo time montado pelo então presidente e pela excessiva troca de técnicos no clube.

Para ilustrar que o efeito da troca de técnico é, até certo ponto, temporária, ou seja, não resolve o problema do rebaixamento, segue a tabela 6.6, com o número de técnicos que cada equipe teve ao longo de todo o campeonato.

Tabela 6.6: Número de Técnicos que cada Equipe Teve no Campeonato de 2008

	Número de Técnicos
São Paulo	1
Grêmio	1
Cruzeiro	1
Palmeiras	1
Flamengo	1
Internacional	2
Botafogo	3
Coritiba	1
Goiás	2
Sport	1

	Número de Técnicos
Vitória	1
Atlético_MG	3
Atlético_PR	5
Fluminense	3
Santos	3
Figueirense	5
Náutico	4
Vasco	3
Portuguesa	3
Ipatinga	3

Com a tabela 6.6, percebe-se que a troca de técnicos não é a solução para salvar a equipe do rebaixamento ou para conseguir a vaga na copa Libertadores da América, pois as 5 primeiras equipes do campeonato não trocaram de técnico durante a competição, enquanto que todos os rebaixados acabaram tendo pelo menos 3 técnicos durante toda a competição.

Muitas vezes a troca de treinador influencia positivamente na equipe, pois segundo Samulski (2002), a mudança na forma de trabalho pode gerar novo ânimo para um grupo de pessoas e isto pode interferir de forma positiva no resultado final do trabalho. Outro ponto que pode ser tratado de forma diferenciada com a mudança de técnico são as situações de jogo que geram *stress* para o time como em De Rose Júnior (2004).

Porém, estas mudanças surtem efeitos imediatos e são creditadas ao novo técnico, quando na verdade, este apenas deu uma nova motivação à equipe. O que pode ser constatado com uma frase de Geninho na entrevista coletiva após a sua estréia vitoriosa na equipe do Goiás quando disse: “Eu só dei dois treinos. Ficaria impossível fazer alguma grande mudança tática na equipe. Tratei apenas de recuperar psicologicamente o Goiás. Técnico que estréia não é milagroso, só estimula o milagre”. Além desta frase do Geninho, o dirigente do Palmeiras, Salvador Hugo Palaia, afirma que a troca de técnico é apenas mais um aspecto motivacional, como na declaração dada após a demissão de mais um técnico, onde ele diz

“Nós temos de estimular o grupo. Não há maneira mais rápida e eficiente do que trocar de comandante. É mais prático e viável do que mandar embora toda a equipe”.

Após as análises dos modelos, serão feitos os testes de validação para o modelo. Serão calculados os testes de razão de verossimilhança e o pseudo-coeficiente de determinação, a partir dos  $\xi(0)$  e o  $\xi(\beta)$ . Na tabela 6.7 temos os dois testes para todos os modelos a partir da rodada 10 até a rodada 38.

Tabela 6.7: Testes – Modelo 1

Rodada	$\xi(0)$	$\xi(\beta)$	Razão de Verossimilhança	Pseudo-Coeficiente de Determinação
10	190.6	158.3	0.831	0.169
11	211.5	178.0	0.842	0.158
12	228.4	194.9	0.853	0.147
13	243.8	217.0	0.890	0.110
14	263.3	235.4	0.894	0.106
15	287.1	259.4	0.904	0.096
16	305.0	275.2	0.902	0.098
17	331.1	294.0	0.888	0.112
18	349.2	310.7	0.890	0.110
19	364.1	324.0	0.890	0.110
20	378.9	336.4	0.888	0.112
21	399.9	357.7	0.895	0.105
22	422.1	378.4	0.897	0.103
23	445.6	404.8	0.908	0.092
24	469.5	426.2	0.908	0.092
25	489.8	450.9	0.920	0.080
26	508.9	466.4	0.916	0.084
27	528.6	488.4	0.924	0.076
28	548.3	501.3	0.914	0.086
29	570.8	522.6	0.916	0.084
30	596.8	549.8	0.921	0.079
31	616.0	570.3	0.926	0.074
32	638.1	590.9	0.926	0.074
33	657.4	615.0	0.936	0.064
34	681.2	637.2	0.935	0.065
35	693.2	649.3	0.937	0.063
36	712.3	667.4	0.937	0.063
37	738.3	695.0	0.941	0.059
38	759.5	716.5	0.943	0.057

Ao observar a tabela, percebe-se que o modelo vai melhorando conforme as rodadas vão passando. Na rodada 10, a estatística do teste apresenta um valor de 0.169, que é completamente aceitável, pois, segundo MacFadden, modelos com 0.200 de pseudo-

coeficiente de determinação podem ser considerado bom. Esses valores vão reduzindo conforme as rodadas aumentam.

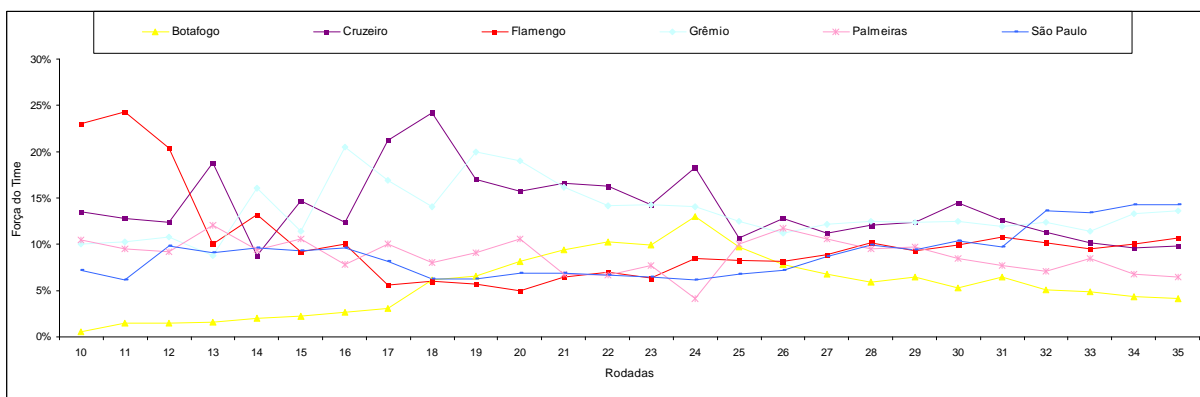
Uma última análise que pode ser percebida pelas tabelas 6.4 e 6.5 é a demora do modelo em identificar as mudanças mais sutis, pois caso o time comece a ter um desempenho ruim, este modelo tende a tardar na captação dessa mudança de comportamento. Um exemplo prático dessa demora é o Flamengo que, entre as rodadas 11 e 18 conseguiu apenas 5 pontos em 24 disputados e mesmo assim continuava sendo a equipe de maior força até a décima quarta rodada, quando foi ultrapassado pela equipe do Grêmio, porém com um atraso na detecção da queda de rendimento do Flamengo.

### 6.2.2 – Modelo 2 - Força do Time, Vantagem de Jogar em Casa e Desempenho Recente Dentro e Fora de Casa

Agora serão efetuadas as análises relacionadas ao segundo modelo desenvolvido, este modelo adiciona o desempenho recente do time mandante jogando em casa e o desempenho do time visitante nos 3 últimos jogos atuando fora de casa. Com estas informações espera-se que as mudanças sejam captadas de forma mais rápida do que no modelo 1, pois como analisado anteriormente, o modelo 1, por considerar todos os jogos com a mesma importância, demora em captar mudanças de rendimento das equipes.

Assim como no primeiro modelo, serão avaliados a força de cada equipe, as chances do time figurar entre os 4 de maior pontuação, assim como as chances do time terminar a competição entre os 4 últimos. A figura 6.8 mostra a força de cada equipe ao utilizar o modelo 2.

Figura 6.8 – Força dos Times - Modelo 2



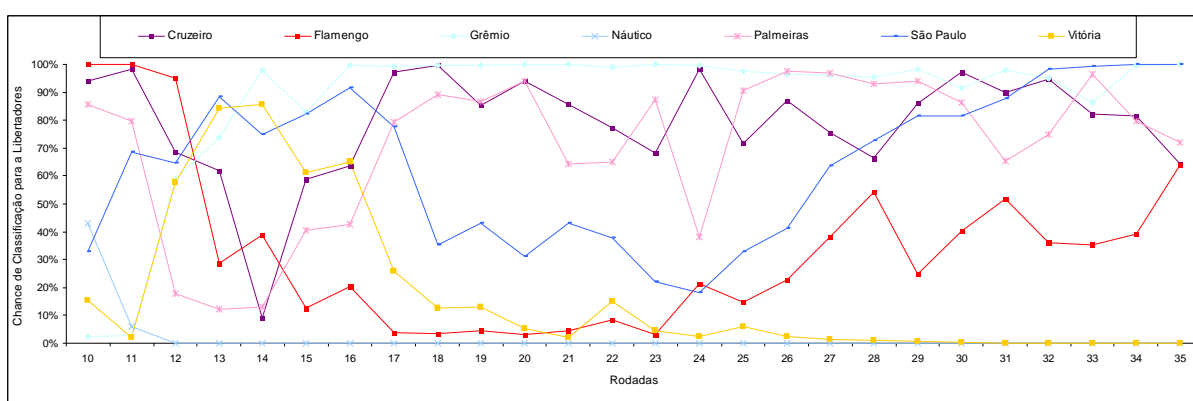
Ao observar a figura 6.8, percebe-se que o Náutico teve uma boa queda na força, pois, como aconteceu no primeiro modelo, esse foi um dos times que tinha bastante força na décima rodada, porém teve a queda de rendimento assim como o Flamengo.

Neste modelo, a força do São Paulo não fica tão explícita, isso porque, apesar de ter feito um segundo turno praticamente sem derrotas, acabou obtendo muitos empates, o que não reflete de forma tão clara a força da equipe.

Um time que conseguiu uma boa melhora foi o Botafogo, pois passou a conseguir um bom desempenho a partir da décima sétima rodada, principalmente, porque ganhou todos os jogos realizados em casa da décima segunda até a vigésima rodada, além de conseguir ganhar dois jogos onde não tinha o mando de campo, por isso conseguiu ter mais força a partir desta rodada.

Agora serão avaliadas as chances de cada equipe figurar entre os 4 de maior pontuação, isto será apresentado no figura 6.9.

Figura 6.9 – Chances de Classificação para a Copa Libertadores da América - Modelo 2



Ao observar a figura 6.9, percebe-se que neste modelo existe uma maior oscilação nas chances de alguns times conseguirem a vaga para Libertadores. Isto se explica, porque o modelo considera o desempenho recente da equipe no modelo, ou seja, para o time mandante os 3 últimos realizados em casa e para o visitante os 3 últimos em que participaram na cada do adversário.

Vale destacar a grande oscilação é a do Palmeiras entre a décima primeira e décima sétima rodada, isto porque, neste período, a equipe realizou boas partidas no Parque Antarctica, onde manda seus jogos, porém, não conseguia pontuar fora de casa.

As equipes do Vitória e Náutico foram as que mais caíram de rendimento no campeonato, tanto que neste modelo tinham boas chances de conseguir uma vaga na

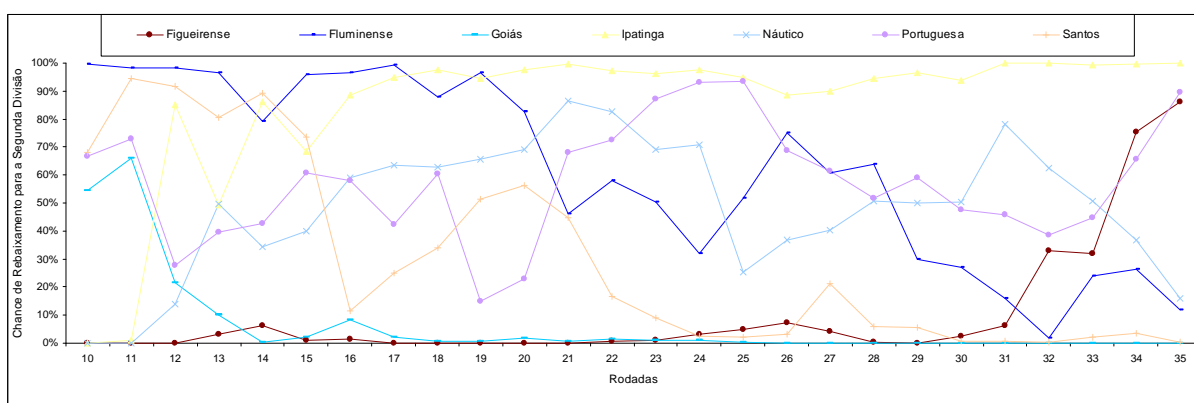


Libertadores no início da competição, porém passaram a não almejar a disputa desta vaga na vigésima quinta para o Vitoria e na décima terceira para o Náutico.

São Paulo também oscilou bastante até a rodada 25, em relação a chance classificação para a competição continental, pois em alguns momentos do campeonato o time não perdia as partidas, mas acabava apenas empatando, o que acaba influenciando nas chances de classificação para a Libertadores.

Na figura 6.10 serão apresentadas as chances de rebaixamento a cada rodada.

Figura 6.10 – Chances de Rebaixamento para a Segunda Divisão - Modelo 2



Assim como no modelo 1, o modelo 2 destaca 3 times que começaram mal a competição cada um com seus motivos, porém foram conseguindo se recuperar ao longo das rodadas e, no final, se salvaram do rebaixamento. São as equipes do Santos, Fluminense e Goiás.

Enquanto estas equipes se recuperaram, o Ipatinga, a partir da rodada 11, não consegue reduzir as chances de rebaixamento, pois foi o primeiro campeonato da série A que este time disputou, por isso, sofreu bastante com a falta de infra-estrutura para disputar uma competição tão longa e difícil.

Vale ressaltar a participação da Portuguesa entre a rodada 17 e 20, com uma recuperação, tanto que sua chance de rebaixamento diminuiu para pouco menos de 20%, porém esta recuperação não durou muito tempo e a equipe voltou ao Z-4 e terminou rebaixada.

Por fim, o Figueirense, que, a partir da vigésima quinta rodada começa a sofrer derrotas em sua casa, que era um dos fatores mais importantes para que o time conseguisse se manter alheio ao rebaixamento. Esse resultado começa a ser captado a partir da rodada 30, com a entrada do time no grupo dos que tinham chance de rebaixamento. Estas derrotas levaram o time a zona do rebaixamento e com remotas chances de se salvar.

De forma análoga foram feitos os mesmos testes estatístico para o modelo 2 e estes estão apresentados na tabela 6.11.

Tabela 6.11: Testes – Modelo 2

Rodada	$\xi(0)$	$\xi(\beta)$	Razão de Verossimilhança	Pseudo-Coeficiente de Determinação
10	190.6	145.2	0.762	0.238
11	211.5	166.8	0.788	0.212
12	228.4	179.0	0.784	0.216
13	243.8	199.6	0.819	0.181
14	263.3	216.5	0.822	0.178
15	287.1	240.7	0.839	0.161
16	305.0	257.8	0.845	0.155
17	331.1	275.7	0.832	0.168
18	349.2	294.0	0.842	0.158
19	364.1	306.8	0.843	0.157
20	378.9	318.1	0.840	0.160
21	399.9	341.1	0.853	0.147
22	422.1	360.7	0.855	0.145
23	445.6	385.2	0.865	0.135
24	469.5	406.6	0.866	0.134
25	489.8	436.1	0.890	0.110
26	508.9	449.1	0.882	0.118
27	528.6	471.1	0.891	0.109
28	548.3	482.5	0.880	0.120
29	570.8	505.4	0.886	0.114
30	596.8	532.8	0.893	0.107
31	616.0	553.3	0.898	0.102
32	638.1	575.0	0.901	0.099
33	657.4	601.0	0.914	0.086
34	681.2	618.8	0.908	0.092
35	693.2	630.9	0.910	0.090
36	712.3	647.2	0.909	0.091
37	738.3	674.5	0.914	0.086
38	759.5	693.2	0.913	0.087

Na tabela 6.11 percebe-se que as rodadas iniciais estão com os testes um pouco acima do aceitável. Com o avanço do campeonato, o modelo vai melhorando e consegue atingir um nível razoável a partir da décima terceira rodada.

Ao comparar o resultado do modelo 2 com o modelo 1, percebe-se uma pequena melhora no acompanhamento dos nuances que um time sofre durante a competição, porém ainda existe uma certa volatilidade no desempenho das equipes. Por exemplo, a equipe do Internacional, que teve campanha muito boa em casa, tem uma força exagerada quando joga em casa e sua força diminuída de forma bastante significativa quando joga fora de seus domínios. Esse fato também gera uma grande oscilação na força das equipes.

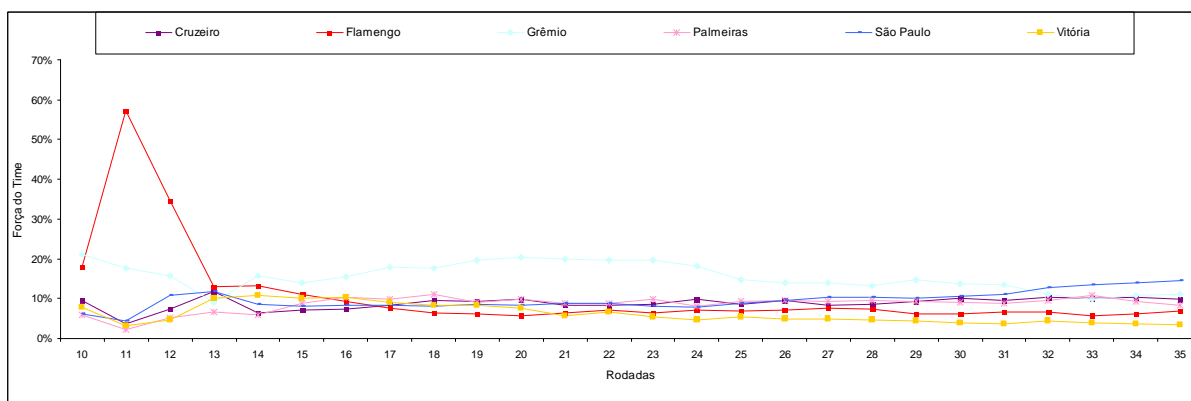
### 6.2.3 – Modelo 3 - Força do Time, Vantagem de Jogar em Casa e Desempenho Recente do Time

Como foi exposto anteriormente, o modelo 2 tem uma melhora na performance recente da equipe em relação ao modelo 1, porém, ainda sofre com equipes que têm um desempenho muito bom em seus domínios e muito ruim fora de casa. Esta oscilação na força das equipes pode ser notada em times como Grêmio e Cruzeiro, que durante algumas rodadas do campeonato conseguiam os pontos em seus domínios, porém, quando eram visitantes, acabavam sendo derrotados e isto influenciava na força do time quando ele atua fora de casa, pois considerava apenas os jogos deste time longe de seus domínios. Agora no modelo 3 será modificada a forma como será utilizada o desempenho recente

A diferença entre o modelo 3 e o modelo 2 é a maneira que a informação do desempenho recente é tratada no modelo. O modelo 3 considera os 6 últimos de cada equipe independente se foi realizado dentro ou fora de casa. Este modelo visa agregar a informação do desempenho recente nas simulações de resultados.

Na figura 6.12 serão apresentadas as forças de cada time com o passar das rodadas do campeonato.

Figura 6.12 – Força dos Times - Modelo 3



Ao analisar esta figura, percebe-se que o São Paulo, Cruzeiro, Grêmio e Palmeiras, foram times que se mantiveram constante durante toda a competição, pois estes times em nenhum momento caíram de rendimento durante o campeonato.

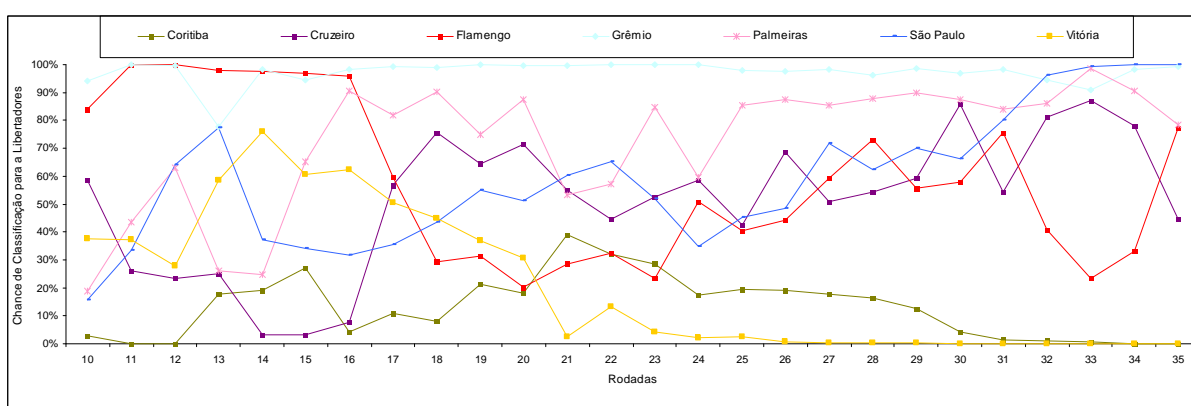
Já o Flamengo e o Vitoria foram 2 times que tiveram seu desempenho afetado, o primeiro em função da saída de vários jogadores para o exterior e o segundo pela falta de um elenco mais forte, pois, com o passar do campeonato, é preciso ter um bom elenco

principalmente por conta de contusões e suspensões as quais o time fica sujeito ao longo da competição.

O Sport foi um time que também teve um aumento de rendimento, principalmente após a conquista da Copa do Brasil, que foi mais uma motivação. Essa conquista já garantia a vaga na Libertadores do ano seguinte. Desta forma, o principal objetivo do time era lutar contra o rebaixamento, o que foi bastante tranquilo, visto que o time tinha um bom elenco.

A partir de agora será analisada as chances de classificação de alguns times para a Libertadores da America. Isto pode ser visto na figura 6.10.

Figura 6.13 – Chances de Classificação para a Copa Libertadores da América - Modelo 3



Este modelo teve uma menor oscilação em relação às chances de classificação para a Libertadores; porque usa um número maior de rodadas, logo o modelo tem um controle maior sobre as modificações ao contrário do modelo 2, onde só os jogos do time da casa como mandantes eram considerados e os jogos dos visitantes como tal eram considerados para o modelo; o que causava uma grande oscilação em relação as chances de classificação para Libertadores.

O Flamengo teve uma queda no rendimento, porém, neste modelo, essa queda só começou a aparecer de forma mais contundente a partir da décima oitava rodada, principalmente porque este time fez apenas 5 pontos da décima até a décima oitava rodada.

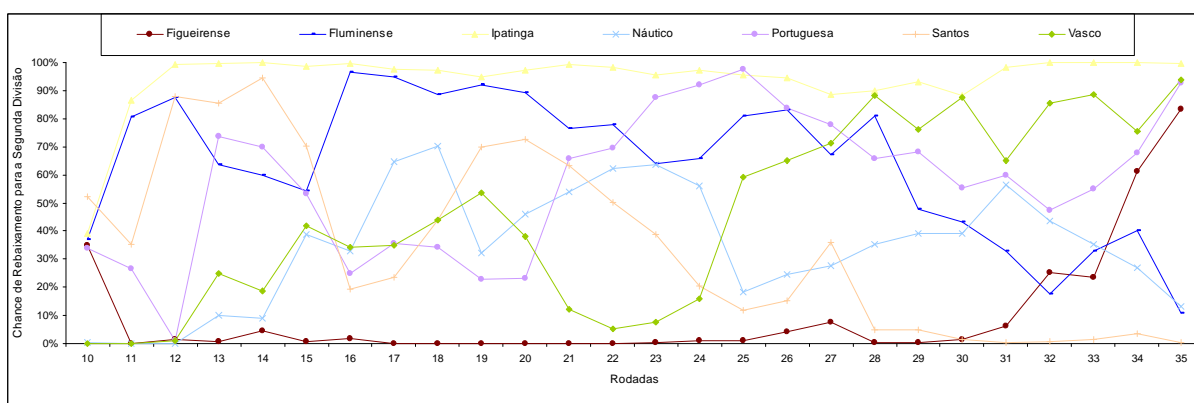
Já o São Paulo, como mencionado nos outros modelos, iniciou a recuperação após encerrar a sua participação na Libertadores e a partir da vigésima rodada o São Paulo pontuou em todas as rodadas do campeonato.

Neste modelo o Botafogo conseguiu ter boas chances de classificação para a Libertadores entre a vigésima primeira e vigésima sexta, isso se deve ao fato de o Botafogo ter tido 100% de aproveitamento nas 5 rodadas anteriores, crescendo bastante na tabela de classificação.

Assim como nos modelos anteriores, vale destacar a constância na campanha das equipes de Palmeiras, Cruzeiro e Grêmio ao longo da competição. Vale ressaltar também o Vitória e Coritiba, que oscilaram durante toda a competição e acabaram em posições intermediárias. O Internacional foi um time que em momento nenhum obteve destaque e boas chances de classificação para a Copa Libertadores por causa do seu desempenho fora de casa, visto que ganhava em casa e perdia fora, desta forma, atingindo somente metade dos pontos que disputava o que o atrapalhou em suas chances de Libertadores.

Em relação ao rebaixamento deste modelo, a figura 6.14 expressa as chances de cada time ao longo da competição.

Figura 6.14 – Chances de Rebaixamento para a Segunda Divisão - Modelo 3



Na figura 6.14, percebe-se que durante todo o campeonato teve uma grande briga contra o rebaixamento, pois pelo menos 8 times estavam disputando para fugir destas que são as posições mais incômodas do campeonato.

A Portuguesa foi um time que esteve durante quase toda a competição na zona do rebaixamento, apesar de ter tido dois momentos de recuperação ao longo do campeonato, com chances de escapar do rebaixamento, a equipe termina o campeonato sendo rebaixada.

Além da Portuguesa, Ipatinga e Vasco foram às outras equipes rebaixadas e que não conseguiram ter uma reação durante a competição, tanto que na trigésima quarta rodada, as duas equipes estavam com pelo menos 90% de chance de serem rebaixados.

Vale ressaltar a queda de rendimento abrupta que o Figueirense sofreu, tanto que na maior parte do campeonato o time não flertou com a zona de rebaixamento, porém, a partir da rodada 30, começa a fazer parte dos times em perigo de serem rebaixados. Ao final da competição, a equipe precisava desesperadamente de vitórias para se salvar, fato este que conseguiu, porém, precisava também de uma combinação de resultados para se salvar, fato este que não ocorreu.

Ao contrario do time do Figueirense o Santos ficou durante todo o primeiro turno com boas chances de rebaixamento, porém, fez um bom segundo turno, escapando assim do rebaixamento.

De forma análoga foram feitos os mesmos testes estatístico para o modelo 3 e estes estão apresentados na tabela 6.15.

Tabela 6.15: Testes – Modelo 3

Rodada	$\xi(0)$	$\xi(\beta)$	Razão de Verossimilhança	Pseudo-Coeficiente de Determinação
10	190.6	129.4	0.679	0.321
11	211.5	143.6	0.679	0.321
12	228.4	162.5	0.712	0.288
13	243.8	191.2	0.784	0.216
14	263.3	209.3	0.795	0.205
15	287.1	228.1	0.795	0.205
16	305.0	249.0	0.817	0.183
17	331.1	269.3	0.813	0.187
18	349.2	280.9	0.804	0.196
19	364.1	294.0	0.807	0.193
20	378.9	302.2	0.798	0.202
21	399.9	324.6	0.812	0.188
22	422.1	346.2	0.820	0.180
23	445.6	368.8	0.828	0.172
24	469.5	387.3	0.825	0.175
25	489.8	408.1	0.833	0.167
26	508.9	423.9	0.833	0.167
27	528.6	446.3	0.844	0.156
28	548.3	468.2	0.854	0.146
29	570.8	493.7	0.865	0.135
30	596.8	521.6	0.874	0.126
31	616.0	540.4	0.877	0.123
32	638.1	564.4	0.884	0.116
33	657.4	583.9	0.888	0.112
34	681.2	606.7	0.891	0.109
35	693.2	619.1	0.893	0.107
36	712.3	638.2	0.896	0.104
37	738.3	669.3	0.907	0.093
38	759.5	688.8	0.907	0.093

Na tabela 6.15 percebe-se que as rodadas iniciais estão com os testes um pouco acima do aceitável, entretanto, assim como no modelo anterior, há uma melhora com o avançar das rodadas, fazendo com que a estatística de aceitação do modelo passe a ser validada a partir da décima sexta rodada.

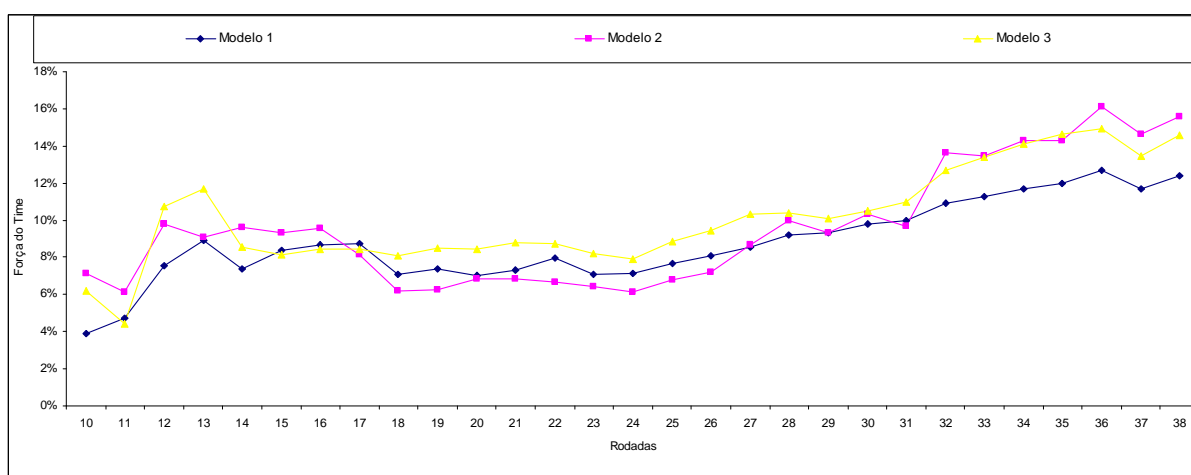
Em relação aos modelos anteriores, percebe-se que houve uma melhora na captação do desempenho recente, tanto que o Flamengo, que era um time que iniciava com a força extremamente alta no modelo, passa a ter uma força mais condizente com seu desempenho no

campeonato e esta mesma equipe vai tendo sua força diminuída com o decorrer da competição, pois passou por um momento muito ruim.

#### 6.2.4 – Comparação dos modelos

Para finalizar as análises dos resultados do modelo, será feita uma pequena comparação dos 3 modelos utilizados, onde serão analisados os efeitos de cada modelo e os possíveis nuances que cada um sofre. Essa análise foi feita para uma equipe somente, que foi a campeã da competição, como ilustrado na figura abaixo:

Figura 6.16 – Força do São Paulo nos três modelos



Em relação à força dos times, os 3 modelos apresentam pequenas diferenças, ou seja, a força do time foi bem captada pelos 3 modelos. O modelo 2 é o que apresenta as maiores oscilações em relação aos demais em função das oscilações agregadas ao modelo a partir da inclusão do desempenho recente, que foi feito com base em 3 jogos.

No que tange às previsões dos modelos para a copa Libertadores da América e para o rebaixamento, o modelo 1 é o modelo mais rígido, ou seja, aquele que demora mais a identificar as mudanças que ocorrem durante a competição, tanto que mantém o Flamengo durante quatro ou cinco rodadas com grandes chances de classificação para Libertadores, enquanto que o Flamengo já sofria com as saídas de jogadores e o baixo rendimento da equipe.

Os modelos 2 e 3 utilizavam, além da força do time e da vantagem de jogar em casa, uma informação de desempenho recente do time.

Para o modelo 2 foi utilizado o momento da equipe conforme o mando de campo, ou seja, se o time jogava em casa, utilizava-se como variável momento os 3 últimos jogos

realizados em casa por esta equipe e se ela era visitante os 3 jogos realizados como tal. Este modelo acabou por oscilar bastante, principalmente pela característica do campeonato, onde vitórias de visitante é um resultado que ocorre uma vez a cada cinco jogos. Em função disso, este modelo sofre bastante com as diferenças de rendimento das equipes quando atuam dentro e fora de casa. A única equipe que sofreu menos em relação às oscilações deste, foi o São Paulo no segundo turno, por ter terminado este turno praticamente invicto.

Por fim o terceiro modelo que foi um meio termo entre o modelo 1 e 2, pois ele considera todos os 6 últimos jogos de cada equipe na hora de considerar a variável momento, por isso ele foi o que conseguiu acompanhar de forma bem fiel as mudanças e efeitos que cada equipe sofreu durante a competição. Ao contrário dos modelos anteriores, o modelo 3 identificou a melhora no rendimento do Botafogo da vigésima até a vigésima quinta rodada, o crescimento do São Paulo ao longo da competição e a queda de Náutico e Figueirense.

Em relação aos 3 modelos apresentados, pode-se concluir que o modelo 3 é mais indicado, visto que, com ele, o desempenho recente das equipes é melhor tratado, pois ele leva em consideração a pontuação obtida pela equipe nos últimos 6 jogos, tratando assim o momento vivenciado pela equipe. Este efeito é bastante importante, pois com ele percebe-se um bom desempenho do Botafogo no meio do campeonato, onde esta equipe chegou a ganhar 5 jogos seguidos. Além deste, o São Paulo foi outra equipe que teve um novo desempenho no modelo, pois cresceu mais rapidamente sua força após a eliminação na Copa Libertadores.



## **7. CONCLUSÕES**

O presente estudo teve objetivo de organizar e desenvolver um modelo que conseguisse prever os resultados do campeonato brasileiro, para isso foram desenvolvidos 3 modelos que abordaram a vantagem de jogar em casa (*home advantage*), a força do time e momento da equipe (trabalhado de duas maneiras diferentes).

Com relação aos modelos, percebeu-se que o modelo 1 era o mais rígido do que os outros por não considerar o momento que o time estava vivendo, ou seja, o modelo trabalhava apenas com o *home advantage* e a força da equipe.

No modelo 2 foi inserida uma variável com o objetivo de tentar verificar o momento, que é uma variável importante, pois existe um ditado que diz que ‘futebol é momento’. Quando esta variável foi inserida no modelo, ocorreu uma melhora nas previsões e no cálculo das chances de Libertadores e rebaixamento.

Porém durante as análises do modelo 2, percebeu-se que existia uma oscilação grande na previsão de resultados, pois o efeito momento foi inserida através do desempenho recente dos times mandante com os 3 últimos jogos que este time realizou em seus domínios, enquanto o visitante teria o seu desempenho recente computando os 3 últimos jogos realizado longe de casa.

Para tentar corrigir este efeito, foi desenvolvido o modelo 3, que considera os últimos 6 jogos independente do mando de campo.

Para avaliar as equipes em relação as chances de rebaixamento e classificação para Libertadores foi desenvolvido o modelo que utiliza os coeficientes para determinar as probabilidade de cada resultado em cada uma das partidas restantes do campeonato. Em seguida, foram gerados números aleatórios que são comparados as probabilidades para cada

partida e assim o resultado da partida é definido. Este método de previsão de resultados se mostrou bastante eficaz.

Em relação aos times, vale destacar o São Paulo que teve um desempenho sensacional no segundo turno do campeonato e sagrou-se campeão, além da regularidade (característica fundamental nesse modelo de campeonato e confirmada a partir do resultado dos jogos) de alguns times como Cruzeiro e Grêmio que como prêmio, encerraram o campeonato classificados para a copa Libertadores da América.

Pode-se destacar alguns rompantes de bom futebol como Botafogo, Vitória, Goiás e Coritiba, que foram times que em 4 ou 5 rodadas seguidas conseguiram fazer boa campanha e com isso entraram num grupo que não estavam com risco de cair para a segunda divisão, mas também não almejavam nenhum objetivo maior no campeonato.

Dentre os rebaixados vale falar sobre o Figueirense que durante boa parte do campeonato fez uma boa campanha (principalmente em casa), porém, da vigésima terceira rodada em diante começou a ter maus resultados em seus domínios e passou a ser um candidato a rebaixamento. Fato este que se consumou ao final do campeonato, mesmo com o time vencendo os últimos 4 jogos.

Os outros 3 times que foram rebaixados: Vasco, Portuguesa e Ipatinga não tiveram um bom desempenho durante toda a competição, o primeiro por estar em um momento muito turbulento no clube, os demais por falta de estrutura e pela falta de um bom elenco para disputar um campeonato tão longo, difícil e disputado como o campeonato brasileiro.

Para trabalhos futuros, poderiam ser inseridas variáveis que identificassem melhor o mando de campo, pois neste estudo foi utilizado o mando de campo determinado pela tabela que foi confeccionada pela Confederação Brasileira de Futebol (CBF), o que nem sempre é real, pois em clássicos regionais o mando de campo nem sempre pode ser considerado, ou então em casos onde uma equipe manda fora de seus domínios por algum tipo de punição.

Pode-se tentar utilizar outras variáveis independentes com o objetivo de melhorar a definição do modelo, tais variáveis poderiam ser a presença de um jogador chave para a equipe ou qualquer outra que possa agregar algum tipo de informação para o modelo.

Além do que foi exposto, pode-se utilizar variáveis nebulosas (lógica fuzzy), assim como uma abordagem com base em meta-heurística.

## **8. REFERÊNCIAS BIBLIOGRÁFICAS**

**ANDERSSON, P.; EDMAN, J.; EKMAN, M.** Predicting the World Cup 2002 in soccer: Performance and confidence of experts and non-experts. *International Journal of Forecasting*, Volume 21, Issue 3, July-September 2005, Pages 565-576, 2005.

**BALMER, N. J.; NEVILL, A. M. ;WILLIAMS, A. M.** Home advantage in the Winter Olympics (1908–1998). *J. Sports Sci.*, 19, 129–139, 2001.

**BALMER, N. J.; NEVILL, A. M. ;WILLIAMS, A. M.** Modelling home advantage in the Summer Olympic Games. *J. Sports Sci.*, 21, 469–478, 2003.

**BEN-AKIVA, M. & LERMAN, S.R.** *Discrete choice analysis: Theory and application to travel demand*, Cambridge: MIT Press, p.59-99,155-215 e 422-453.1985.

**BRISKORN, D. ; DREXL, A; SPIEKSM, F.C.R.** Round robin tournaments and three index assignment, *Manuskripte aus den Instituten faur Betriebswirtschaftslehre der Universitaat Kiel*, 2006.

**CALÔBA, G. M., LINS, M.P.E.,** *Performance Assessment of the soccer teams in Brazil using DEA*. *Pesquisa Operacional*, v.26, n.3,p.521-536, 2006.

**CORDEIRO, G. M..** *Modelos Lineares Generalizados*. VII SINAPE, Simpósio Nacional de Probabilidade e Estatística. 1986.

**CHENG, T.; CUI, D., FAN; Z.; ZHOU, J.; LU, S.,** *A New Model to Forecast the Results of Matches Based on Hybrid. Neural Networks in the Soccer Rating System*. 5th IEEE International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2003, Oral), pp. 308-313, 2003.

**CHURILOV, L., FLITMAN, A..** *Towards fair ranking of olympics achievements: The case of Sydney 2000*. *Computers and Operations Research*, 33 (7), pp. 2057-2082, 2006.

**COURNEYA, K. S.; CARRON, A. V.** *The home advantage in sport competitions: a literature review.* Journal of Sport and Exercise Psychology; 14:13-27. 1992.

**CRAIG, L. A.; HALL, A. R.** *Trying out for the team: Do exhibitions matter? Evidence from the National Football League.* Journal of the American Statistical Association, 89, 1091-1099. 1994.

**DARE, W. H.; MACDONALD, S.** *A Generalized Model for Testing the Home Advantage and Favorite Team Advantage in Point Spread Markers.* Journal of Financial Economics, 40 (1996), 295-318.

**DE ROSE JUNIOR, D.** Situações de jogo como fontes de “stress” em modalidades esportivas coletivas. Revista Brasileira de Educação Física Esportiva, 18(4):385-395, 2004.

**FREITAS FILHO, P. J.** *Introdução à Modelagem e Simulação de Sistemas*, Visual Books. 2001.

**GELADE, G. A.; DOBSSON, P.** *Predicting the Comparative Strengths of National Football Teams.* Social Science Quarterly, volume 88, 1, March 2007.

**GUIMARÃES, I. A.** *Construção e avaliação de uma regra de reconhecimento e classificação de clientes de uma instituição financeira com base na análise multivariada.* 2000. 87 f. (Dissertação de Mestrado) - Métodos Numéricos em Engenharia, Universidade Federal do Paraná, Curitiba, 2000.

**HAIGH, J.;** *Uses and limitations of mathematics in sport.* IMA Journal of Management Mathematics 20, 97–108, 2009.

**ÍBSAN, ALP;** *Performance Evaluation of Goalkeepers of the World Cup.* Journal of Science 19 (2): 119-125. 2006.

**LAWAL, H. B.; SUNDHEIM, A. R.** *Modeling 2002-2003 English Premier League Results.* IMA Sport 2007, 115-124. 2007.

**LEE, A. J.** *Modeling scores in the premier league: Is Manchester United really the best?* Chance, 10, 15-19. 1997.

**LINS, M.P.E.A , GOMES, E.G.B , SOARES DE MELLO, J.C.C.B., SOARES DE MELLO, A. J. R.** *Olympic ranking based on a zero sum gains DEA model.* European Journal of Operational Research, 148 (2), pp. 312-322. 2003.

**LOUVIERE, J. J. ; HENSHER, D. A; SWAIT, J. D.,** *Stated Choice Methods – Analysis and Application.* 1ª ed. Cambridge, Cambridge University Press, 2000.

- MARTINS, C. A. C. (1988).** *Análise de regressão logística*. 53 f. (Dissertação de Mestrado) – Centro de Ciências Exatas e da Natureza, Universidade Federal de Pernambuco, Recife, 1988.
- McFADDEN, D.** *Conditional logit analysis of qualitative choice behavior*. In *frontiers in Econometric*. Nova Iorque; Academic Press, ed. P. Zarembka, 105-142, 1974.
- MCHALE, I., SCARF, P..** *Forecasting international soccer match results using bivariate discrete distributions*. Salford Business School Working Paper Series. Paper no. 322/2006
- NEVILL, A. M.; NEWELL, S. M.; GALE, S..** *Factors Associated With Home Advantage in English and Scottish Soccer Matches*. *Journal of Sports Science*, volume 14, 181-186, April 1996.
- PEGDEN, C.D.; SHANNON, R.E., SADOWSKI, R. P.** *Introduction to Simulation Using SIMAN*, McGraw-Hill, NY, 2<sup>nd</sup> ed.
- POLLARD, R.** *Home advantage in soccer: a retrospective analysis*. *Journal of Sports Sciences*, 4, 237-248. 1986.
- POLLARD, R..** *Evidence of a reduced home advantage when a team moves to a new stadium*. *Journal of Sports Sciences*, 20 (12), pp. 969-973002E. 2002.
- POLLARD, R.; POLLARD G.;** *Long-Term Trends in Home Advantage in Professional Team Sports in North America and England (1876 – 2003)*. *Journal of Sports Sciences*, Volume 23, April 2005, 337-350. 2005.
- POLLARD, R..** *Worldwide regional variations in home advantage in association football*. *Journal of Sports Sciences*, 24 (3), pp. 231-240. 2006,
- ROCHE, F. P.** *Gestão desportiva. Planejamento estratégico nas organizações esportivas*. 2.ed. Porto Alegre: Artmed. 2002.
- ROTZ, W. e E. FALK, D. W., e Mulrow, J.,** *A Comparison of Random Number Generators Used in Business*, presented at Joint Statistical Meetings, Atlanta, GA, 2001
- RUGGIERO, M.A.G. ; LOPES, V. L.R. .** *Cálculo Numérico: Aspectos Teóricos e Computacionais*. Ed. Makron Books, 2ª Edição, São Paulo, 1996.
- SAMULSKI, D.M.** *Psicologia no esporte*. São Paulo, Manole, 2002.
- SANT’ANNA, A. P.; BARBOZA, E. U.; SOARES DE MELLO, J.C.C.B..** *A framework to explain the Brazilian soccer championships classifications*. *IMA Sport* 2007, 180-185. 2007,

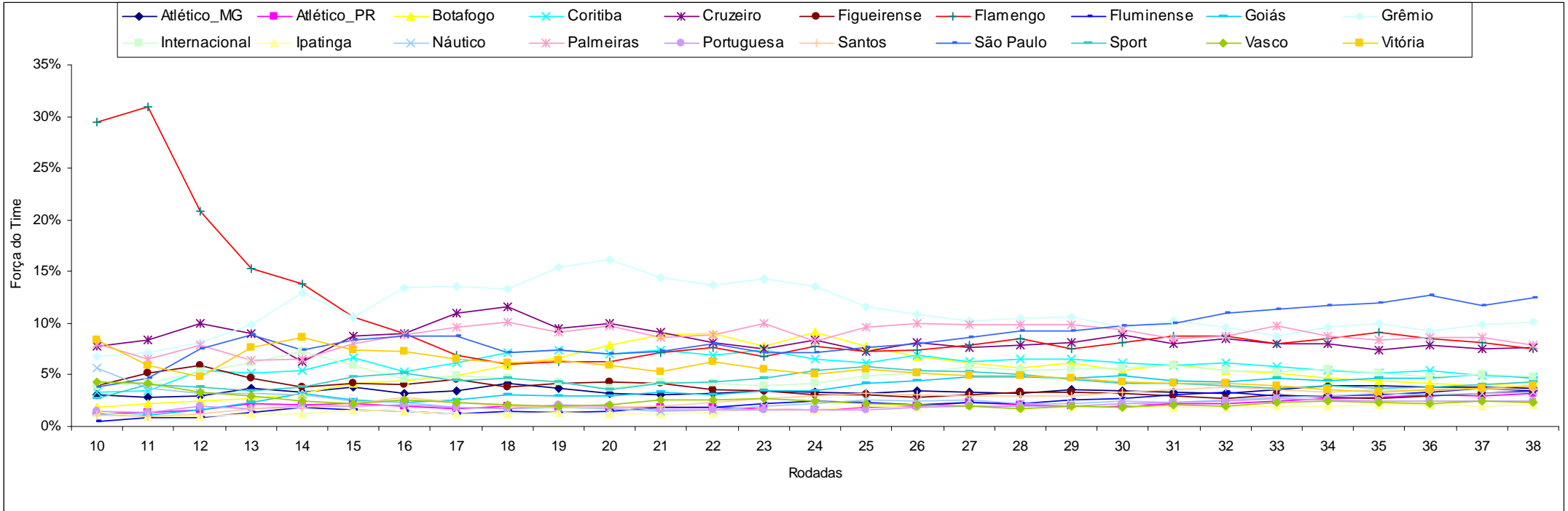
- SANT'ANNA, A. P.; BARBOZA, E. U.; SOARES DE MELLO, J.C.C.B..** *Classification of the teams in the Brazilian Soccer Championship by probabilistic criteria composition.* Soccer and Society. 2010 (no prelo).
- SCARFF, P. A.; SHI X..** *The importance of a match in a tournament.* Computers & Operations Research 35, 2406-2418. 2008.
- SMALL, K; HSIAO, C..** *Multinomial Logit Specifications Tests.* Working Paper. Department of Economics, Princeton University, Princeton, NJ, 1982.
- SPANN, M.; SKIERA, B..** *Sports Forecasting: A Comparison of the Forecast Accuracy of Prediction Markets, Betting Odds and Tipsters.* Journal of Forecasting. 2008,
- STEFANI, R. T. E CLARKE, S. R.,** *Predictions and home advantage for Australian Rules football,* Journal of Applied Statistics, 9, 251-261. 1992.
- URRUTIA. S.; CARDOZO E SOUZA, A., STANCIOLI, P.,RIBEIRO, C.C. .** *Maximização de quebras em torneios Round Robin simples.* XXXIX Simpósio Brasileiro de Pesquisa Operacional, 2007.
- WATSON, H.J. BLACKSTONE JR, J. H..** *Computer simulation,* 2nd ed. New York: John Wiley & Sons. 1989.
- WICHMAN, B.A. e I.D. HILL,** *Building a Random-Number Generator,* BYTE, pp. 127-128, Março 1987.
- YIANNAKIS, A.; SELBY, M.J.P.; DOUVIS, J.; HAN, J.Y..** *The Power of Social Context — A Time Series Analysis with English Premier League Soccer.* International Review for the Sociology of Sport, Vol. 41, No. 1, 89-115 (2006).

## 9. ANEXOS

### ANEXO A – Força dos Times no modelo 1

Modelo 1	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
Atlético_MG	3,1%	2,8%	2,9%	3,7%	3,3%	3,8%	3,2%	3,5%	4,2%	3,7%	3,2%	3,1%	3,2%	3,4%	3,3%	3,3%	3,4%	3,4%	3,2%	3,6%	3,4%	3,4%	3,3%	3,6%	3,9%	4,0%	3,8%	3,8%	3,7%
Atlético_PR	1,3%	1,3%	1,7%	2,3%	2,1%	2,2%	2,0%	1,8%	1,9%	1,8%	1,9%	1,9%	1,8%	1,7%	1,7%	1,9%	2,0%	2,3%	2,1%	2,0%	1,9%	2,2%	2,3%	2,5%	2,8%	2,9%	3,1%	3,0%	3,2%
Botafogo	1,8%	2,2%	2,5%	2,8%	3,3%	4,2%	4,5%	4,9%	6,0%	6,6%	7,8%	8,9%	9,0%	7,8%	9,1%	7,7%	6,8%	6,1%	5,7%	6,1%	5,4%	6,0%	5,5%	5,1%	4,7%	4,5%	4,2%	3,9%	4,3%
Coritiba	3,3%	3,5%	5,4%	5,2%	5,4%	6,7%	5,3%	6,1%	7,1%	7,4%	7,0%	7,4%	6,9%	7,4%	6,5%	6,2%	6,8%	6,3%	6,5%	6,5%	6,2%	5,9%	6,1%	5,8%	5,4%	5,2%	5,4%	5,0%	4,8%
Cruzeiro	7,7%	8,4%	10,0%	9,0%	6,3%	8,7%	9,0%	10,9%	11,6%	9,5%	10,0%	9,2%	8,1%	7,5%	8,4%	7,3%	8,1%	7,6%	7,9%	8,1%	8,9%	8,0%	8,5%	8,0%	8,1%	7,5%	7,9%	7,5%	7,6%
Figueirense	4,0%	5,1%	5,9%	4,7%	3,8%	4,2%	4,0%	4,5%	3,8%	4,2%	4,3%	4,1%	3,6%	3,4%	3,0%	3,1%	2,8%	3,1%	3,4%	3,4%	3,2%	3,0%	2,7%	3,1%	2,8%	2,8%	2,9%	3,2%	3,4%
Flamengo	29,4%	30,9%	20,8%	15,3%	13,8%	10,6%	9,0%	6,9%	6,0%	6,3%	6,2%	7,1%	7,6%	6,8%	7,7%	7,2%	7,4%	7,9%	8,6%	7,5%	8,2%	8,7%	8,7%	8,0%	8,6%	9,1%	8,5%	8,1%	7,5%
Fluminense	0,5%	0,8%	0,9%	1,4%	1,8%	1,6%	1,5%	1,2%	1,5%	1,4%	1,5%	1,9%	1,9%	2,3%	2,4%	2,4%	2,1%	2,3%	2,3%	2,5%	2,8%	3,0%	3,3%	3,0%	2,9%	3,1%	3,4%	3,7%	3,5%
Goiás	1,1%	1,1%	1,6%	2,4%	3,2%	2,6%	2,3%	2,6%	3,0%	2,9%	3,0%	3,4%	3,0%	3,4%	3,5%	4,2%	4,4%	4,8%	4,8%	4,6%	5,0%	4,5%	4,3%	4,7%	4,4%	4,7%	4,7%	5,1%	4,7%
Grêmio	6,8%	7,1%	8,2%	9,9%	12,9%	10,6%	13,5%	13,5%	13,3%	15,4%	16,1%	14,5%	13,7%	14,3%	13,6%	11,5%	10,8%	10,2%	10,4%	10,7%	9,7%	10,2%	9,6%	8,8%	9,6%	10,0%	9,3%	9,9%	10,1%
Internacional	4,0%	4,2%	5,2%	6,2%	7,3%	5,9%	4,5%	4,9%	4,7%	4,3%	3,8%	4,3%	4,1%	4,0%	4,1%	4,9%	5,2%	5,8%	5,4%	5,6%	5,7%	5,9%	5,4%	5,3%	5,5%	5,2%	4,7%	5,1%	4,8%
Ipatinga	1,3%	0,9%	0,9%	1,3%	1,2%	1,6%	1,5%	1,3%	1,3%	1,4%	1,3%	1,2%	1,4%	1,7%	1,6%	1,9%	1,9%	2,0%	1,8%	1,8%	2,0%	1,9%	1,8%	2,0%	2,0%	2,1%	2,1%	2,0%	2,1%
Náutico	5,7%	3,7%	3,2%	3,2%	3,1%	2,5%	2,3%	1,9%	1,7%	1,9%	1,7%	1,5%	1,6%	2,0%	2,2%	2,6%	2,5%	2,6%	2,2%	2,2%	2,5%	2,4%	2,5%	2,8%	3,0%	3,2%	3,0%	3,2%	3,3%
Palmeiras	8,1%	6,6%	7,9%	6,4%	6,5%	8,0%	8,8%	9,6%	10,1%	9,2%	9,7%	8,6%	8,9%	9,9%	8,2%	9,6%	10,0%	9,9%	9,9%	9,9%	9,3%	8,5%	8,8%	9,8%	8,8%	8,4%	8,6%	8,7%	7,9%
Portuguesa	1,5%	1,4%	2,0%	1,8%	1,9%	2,0%	2,1%	1,9%	1,8%	2,1%	1,9%	1,6%	1,6%	1,7%	1,6%	1,6%	1,8%	2,0%	1,9%	2,0%	2,2%	2,3%	2,4%	2,7%	2,6%	2,5%	2,5%	2,5%	2,4%
Santos	1,1%	1,0%	1,2%	1,8%	1,8%	2,2%	2,9%	2,3%	2,0%	1,9%	1,9%	2,0%	2,3%	2,7%	3,0%	3,3%	3,1%	2,9%	3,0%	2,9%	3,4%	3,6%	3,8%	3,6%	3,4%	3,6%	3,5%	3,6%	3,5%
São Paulo	3,9%	4,7%	7,6%	8,9%	7,4%	8,4%	8,7%	8,7%	7,1%	7,4%	7,0%	7,3%	8,0%	7,1%	7,2%	7,7%	8,1%	8,6%	9,2%	9,3%	9,8%	10,0%	10,9%	11,3%	11,7%	12,0%	12,7%	11,7%	12,4%
Sport	2,7%	4,0%	3,9%	3,4%	3,8%	4,8%	5,2%	4,5%	4,6%	4,4%	3,6%	4,1%	4,3%	4,7%	5,4%	5,8%	5,5%	5,3%	5,0%	4,6%	4,2%	4,2%	4,0%	3,7%	3,9%	3,7%	3,9%	4,0%	4,3%
Vasco	4,3%	4,2%	3,4%	2,9%	2,5%	2,2%	2,5%	2,3%	2,0%	1,9%	2,1%	2,6%	2,6%	2,7%	2,4%	2,2%	2,1%	2,0%	1,7%	2,0%	1,8%	2,1%	2,0%	2,3%	2,5%	2,4%	2,2%	2,5%	2,3%
Vitória	8,4%	5,9%	4,8%	7,6%	8,6%	7,3%	7,3%	6,6%	6,2%	6,4%	5,9%	5,3%	6,3%	5,5%	5,0%	5,6%	5,2%	4,9%	4,9%	4,7%	4,4%	4,2%	4,1%	3,9%	3,5%	3,3%	3,6%	3,7%	4,0%

### ANEXO B – Força dos Times no modelo 1

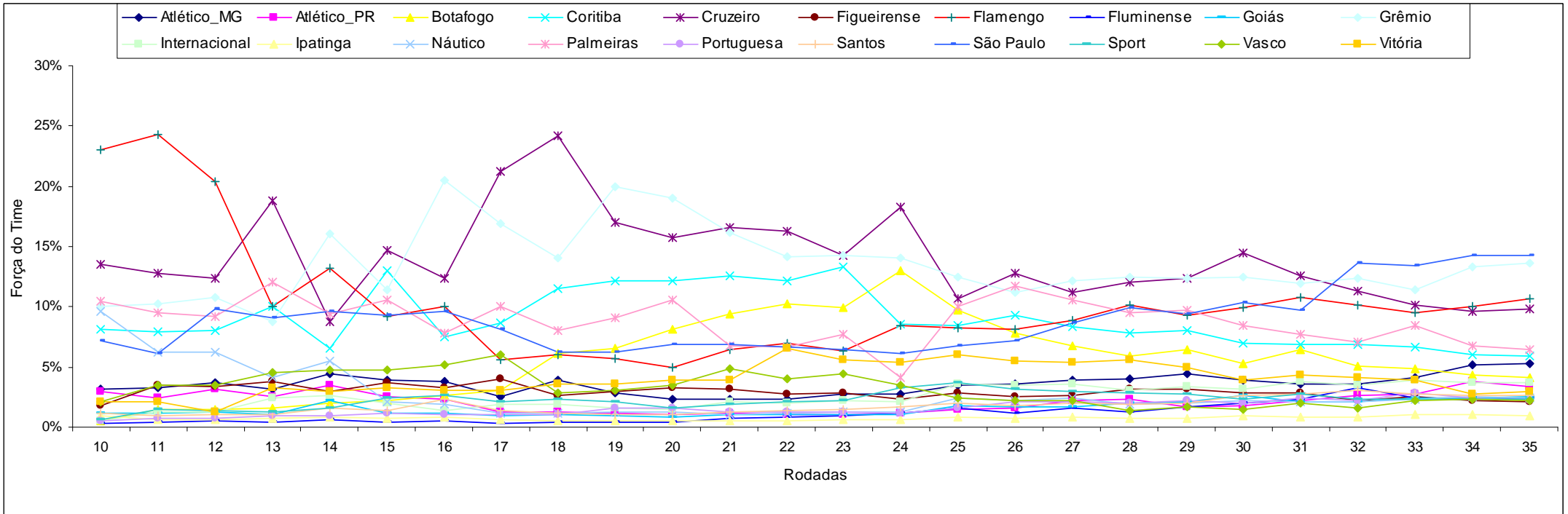




## ANEXO C – Força dos Times no modelo 2

Modelo 2	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
Atlético_MG	3,1%	3,2%	3,7%	3,2%	4,5%	3,9%	3,8%	2,5%	3,9%	2,9%	2,3%	2,4%	2,4%	2,7%	2,8%	3,5%	3,6%	3,9%	4,0%	4,4%	3,9%	3,6%	3,6%	4,1%	5,1%	5,3%	5,3%	5,4%	5,4%
Atlético_PR	3,0%	2,4%	3,2%	2,5%	3,5%	2,5%	2,3%	1,3%	1,3%	1,2%	1,1%	1,2%	1,1%	1,1%	1,1%	1,4%	1,5%	2,2%	2,3%	1,7%	1,8%	2,3%	2,6%	2,8%	3,8%	3,4%	3,8%	3,8%	3,8%
Botafogo	0,6%	1,4%	1,5%	1,6%	2,0%	2,2%	2,7%	3,1%	6,1%	6,5%	8,2%	9,4%	10,2%	9,9%	13,0%	9,7%	7,9%	6,7%	5,9%	6,4%	5,3%	6,4%	5,0%	4,9%	4,3%	4,1%	3,9%	3,6%	3,9%
Coritiba	8,1%	8,0%	8,0%	10,1%	6,6%	13,0%	7,5%	8,7%	11,5%	12,1%	12,1%	12,6%	12,2%	13,4%	8,5%	8,4%	9,3%	8,3%	7,8%	8,0%	7,0%	6,8%	6,8%	6,6%	6,0%	5,9%	5,8%	4,9%	4,9%
Cruzeiro	13,5%	12,8%	12,4%	18,8%	8,7%	14,6%	12,4%	21,2%	24,2%	17,0%	15,7%	16,6%	16,2%	14,3%	18,3%	10,7%	12,7%	11,1%	12,1%	12,3%	14,5%	12,5%	11,3%	10,1%	9,6%	9,9%	10,3%	10,3%	10,0%
Figueirense	1,8%	3,5%	3,4%	3,9%	3,0%	3,7%	3,3%	4,0%	2,7%	3,0%	3,2%	3,2%	2,7%	2,8%	2,4%	2,8%	2,5%	2,6%	3,1%	3,1%	2,9%	2,8%	2,3%	2,6%	2,2%	2,1%	2,2%	2,6%	2,6%
Flamengo	23,0%	24,3%	20,4%	10,1%	13,2%	9,2%	10,1%	5,7%	6,0%	5,7%	5,0%	6,4%	7,0%	6,3%	8,4%	8,2%	8,2%	8,9%	10,1%	9,3%	9,9%	10,8%	10,1%	9,5%	10,0%	10,7%	10,3%	10,1%	9,9%
Fluminense	0,3%	0,4%	0,5%	0,4%	0,7%	0,4%	0,6%	0,3%	0,4%	0,4%	0,4%	0,8%	0,8%	1,0%	1,2%	1,5%	1,2%	1,6%	1,3%	1,7%	2,0%	2,3%	3,2%	2,5%	2,4%	2,4%	3,0%	3,3%	3,1%
Goiás	1,2%	1,2%	1,3%	1,1%	2,2%	1,2%	1,2%	1,0%	1,0%	1,0%	0,8%	1,0%	1,0%	1,1%	1,1%	1,8%	1,7%	1,7%	2,0%	2,1%	2,7%	2,1%	2,1%	2,5%	2,4%	2,5%	2,6%	2,8%	2,5%
Grêmio	10,1%	10,2%	10,8%	8,8%	16,1%	11,4%	20,5%	16,9%	14,0%	20,0%	19,0%	16,2%	14,2%	14,3%	14,0%	12,4%	11,2%	12,1%	12,5%	12,3%	12,4%	11,9%	12,3%	11,4%	13,4%	13,6%	12,0%	12,8%	12,7%
Internacional	0,6%	1,3%	1,3%	2,5%	2,6%	2,0%	1,4%	1,9%	1,9%	1,6%	1,5%	2,1%	2,0%	2,2%	2,1%	3,6%	3,5%	3,6%	3,0%	3,3%	3,2%	3,8%	3,5%	3,9%	3,7%	3,7%	2,6%	2,8%	2,7%
Ipatinga	1,0%	0,6%	0,6%	0,7%	0,8%	0,8%	0,8%	0,6%	0,6%	0,6%	0,5%	0,5%	0,5%	0,6%	0,6%	0,8%	0,8%	0,8%	0,7%	0,7%	1,0%	0,9%	0,8%	1,0%	1,0%	1,0%	1,0%	1,0%	1,1%
Náutico	9,6%	6,3%	6,2%	4,1%	5,5%	2,1%	1,9%	1,1%	1,1%	1,3%	1,2%	1,0%	1,2%	1,3%	1,3%	2,5%	2,2%	2,4%	1,9%	2,0%	2,3%	2,2%	2,1%	2,4%	2,5%	2,4%	2,2%	2,3%	2,4%
Palmeiras	10,5%	9,5%	9,2%	12,0%	9,4%	10,6%	7,9%	10,1%	8,0%	9,1%	10,5%	6,7%	6,6%	7,7%	4,1%	10,0%	11,7%	10,6%	9,5%	9,7%	8,4%	7,7%	7,1%	8,4%	6,7%	6,5%	6,3%	6,7%	5,6%
Portuguesa	0,5%	0,8%	0,7%	0,9%	0,9%	1,2%	1,0%	1,1%	1,1%	1,6%	1,6%	1,3%	1,3%	1,2%	1,2%	1,6%	2,1%	2,1%	2,1%	2,2%	2,1%	2,4%	2,3%	2,8%	2,4%	2,2%	2,2%	2,1%	2,2%
Santos	1,2%	0,9%	1,0%	1,1%	1,6%	1,4%	2,3%	1,4%	1,2%	1,1%	1,0%	1,3%	1,4%	1,5%	1,7%	2,0%	1,8%	2,0%	1,9%	2,0%	2,6%	2,7%	3,1%	2,8%	2,6%	2,6%	2,6%	2,6%	2,5%
São Paulo	7,1%	6,1%	9,8%	9,1%	9,6%	9,3%	9,6%	8,2%	6,2%	6,3%	6,8%	6,8%	6,7%	6,4%	6,1%	6,8%	7,2%	8,7%	10,0%	9,4%	10,3%	9,7%	13,7%	13,4%	14,3%	14,3%	16,1%	14,6%	15,6%
Sport	0,7%	1,4%	1,4%	1,3%	1,6%	2,4%	2,6%	2,1%	2,3%	2,1%	1,6%	1,9%	2,1%	2,2%	3,3%	3,7%	3,2%	3,0%	2,8%	2,7%	2,3%	2,7%	2,3%	2,3%	2,3%	2,3%	2,3%	2,5%	2,5%
Vasco	2,1%	3,5%	3,5%	4,5%	4,7%	4,7%	5,2%	6,0%	2,9%	3,0%	3,5%	4,8%	4,0%	4,4%	3,5%	2,4%	2,3%	2,2%	1,4%	1,7%	1,5%	2,0%	1,6%	2,2%	2,4%	2,2%	1,9%	2,3%	2,2%
Vitória	2,1%	2,1%	1,3%	3,3%	3,0%	3,3%	3,1%	3,0%	3,6%	3,6%	3,9%	3,9%	6,6%	5,6%	5,4%	6,0%	5,5%	5,4%	5,6%	5,0%	4,0%	4,3%	4,1%	3,9%	2,8%	3,0%	3,5%	3,6%	4,2%

ANEXO D – Força dos Times no modelo 2



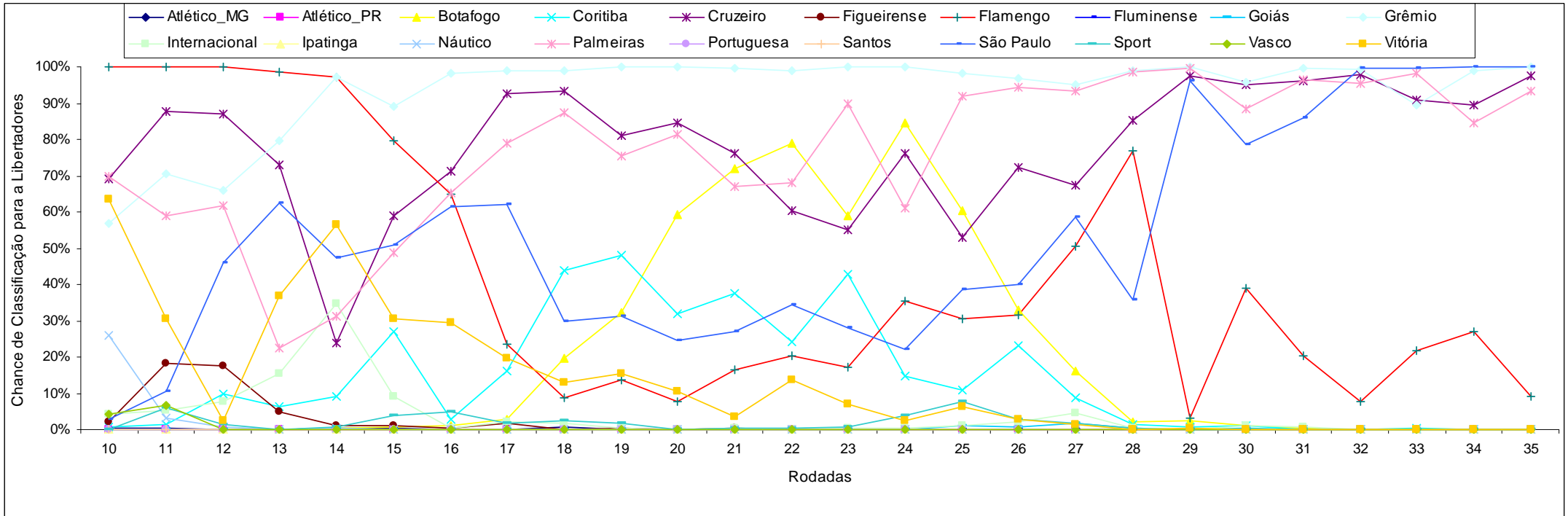
## ANEXO E – Força dos Times no modelo 3

Modelo 3	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
Atlético_MG	1,6%	0,7%	1,4%	2,4%	2,0%	2,1%	1,8%	1,7%	2,3%	2,1%	2,1%	1,8%	2,2%	2,4%	2,5%	2,4%	2,5%	2,9%	2,8%	3,5%	3,3%	3,1%	3,0%	3,3%	3,6%	3,7%	3,5%	3,3%	3,3%
Atlético_PR	1,2%	0,5%	1,3%	2,5%	2,2%	2,3%	2,1%	2,0%	2,2%	2,2%	2,2%	2,0%	2,0%	1,9%	2,2%	2,3%	2,5%	2,2%	1,9%	1,9%	2,0%	2,1%	2,3%	2,6%	2,7%	2,9%	2,9%	3,1%	
Botafogo	1,3%	0,4%	1,5%	2,5%	2,8%	3,4%	3,3%	3,4%	3,6%	4,2%	4,3%	5,7%	6,2%	6,5%	7,6%	7,0%	6,6%	5,7%	5,1%	5,3%	4,9%	5,4%	4,9%	4,5%	4,2%	4,0%	3,9%	3,6%	4,2%
Coritiba	2,3%	0,9%	1,8%	5,2%	5,4%	5,4%	4,2%	5,3%	4,9%	5,7%	5,3%	6,2%	5,6%	5,5%	5,1%	4,9%	5,2%	5,1%	5,5%	5,5%	5,4%	4,9%	5,1%	5,0%	4,5%	4,5%	4,6%	4,3%	4,2%
Cruzeiro	9,6%	3,8%	7,3%	11,9%	6,3%	7,2%	7,3%	8,4%	9,7%	9,3%	9,9%	8,4%	8,4%	8,6%	9,9%	8,6%	9,7%	8,4%	8,5%	9,3%	10,2%	9,7%	10,3%	10,2%	10,2%	9,8%	10,4%	10,0%	10,2%
Figueirense	1,6%	0,8%	2,2%	3,9%	3,1%	4,0%	4,1%	3,7%	3,6%	4,4%	4,8%	4,5%	3,7%	3,6%	3,3%	3,5%	3,1%	2,9%	3,3%	3,3%	3,1%	2,9%	2,7%	3,1%	2,8%	2,7%	2,8%	3,1%	3,3%
Flamengo	17,9%	57,2%	34,6%	12,9%	13,3%	11,0%	9,4%	7,6%	6,5%	6,1%	5,7%	6,4%	7,1%	6,4%	7,0%	6,8%	7,2%	7,5%	7,4%	6,1%	6,2%	6,7%	6,5%	5,8%	6,2%	6,8%	6,4%	6,2%	5,7%
Fluminense	0,2%	0,1%	0,1%	0,3%	0,3%	0,3%	0,5%	0,3%	0,3%	0,3%	0,3%	0,3%	0,7%	0,9%	1,1%	1,1%	0,9%	1,3%	1,4%	1,7%	1,9%	1,9%	2,0%	1,9%	1,9%	2,1%	2,3%	2,8%	2,8%
Goiás	0,5%	0,3%	0,9%	1,8%	2,6%	2,2%	2,0%	2,3%	2,5%	2,4%	2,5%	2,7%	2,4%	2,7%	2,8%	3,5%	3,6%	3,8%	5,0%	5,0%	5,5%	5,2%	5,0%	5,3%	4,8%	5,0%	5,1%	5,2%	4,8%
Grêmio	21,0%	17,8%	15,6%	9,2%	15,8%	14,1%	15,4%	17,9%	17,7%	19,6%	20,3%	19,9%	19,8%	19,7%	18,1%	14,8%	14,0%	14,0%	13,3%	14,9%	13,7%	13,5%	11,1%	9,8%	10,7%	11,0%	10,2%	11,0%	10,9%
Internacional	3,0%	1,2%	3,2%	6,1%	8,1%	6,2%	4,7%	5,0%	4,6%	4,0%	3,8%	4,3%	4,3%	4,0%	4,2%	4,8%	5,2%	5,5%	5,4%	5,4%	5,4%	5,9%	5,3%	5,1%	5,3%	5,0%	4,6%	4,8%	4,7%
Ipatinga	0,9%	0,3%	0,3%	0,9%	0,9%	1,0%	1,0%	1,2%	1,1%	1,2%	1,1%	1,1%	1,1%	1,4%	1,3%	1,5%	1,5%	1,7%	1,6%	1,7%	2,0%	1,8%	1,6%	1,8%	1,8%	1,9%	1,9%	1,8%	1,9%
Náutico	5,8%	2,1%	2,1%	2,2%	2,1%	1,9%	2,0%	1,6%	1,4%	1,6%	1,5%	1,4%	1,5%	1,6%	1,8%	2,1%	2,1%	2,2%	1,8%	1,8%	2,0%	2,0%	2,3%	2,5%	2,7%	2,9%	2,8%	2,9%	3,2%
Palmeiras	6,0%	2,1%	5,1%	6,6%	5,9%	8,9%	10,4%	9,9%	11,0%	9,2%	9,8%	8,9%	8,9%	9,7%	8,1%	9,4%	9,7%	9,3%	9,7%	9,4%	9,1%	8,9%	9,5%	10,7%	9,4%	8,5%	8,7%	8,7%	7,5%
Portuguesa	0,9%	0,4%	1,4%	1,5%	1,5%	1,7%	2,1%	1,6%	1,6%	1,7%	1,6%	1,3%	1,3%	1,4%	1,3%	1,3%	1,3%	1,7%	1,7%	1,7%	1,8%	1,9%	2,2%	2,5%	2,4%	2,2%	2,2%	2,2%	2,2%
Santos	0,8%	0,4%	0,5%	1,4%	1,1%	1,4%	2,2%	1,7%	1,5%	1,4%	1,3%	1,5%	1,6%	2,1%	2,3%	2,5%	2,2%	2,2%	2,4%	2,4%	2,8%	2,9%	2,9%	2,9%	2,7%	2,9%	2,8%	3,0%	2,8%
São Paulo	6,2%	4,4%	10,7%	11,7%	8,5%	8,2%	8,5%	8,4%	8,1%	8,5%	8,4%	8,8%	8,7%	8,2%	7,9%	8,9%	9,5%	10,3%	10,4%	10,1%	10,5%	11,0%	12,7%	13,4%	14,1%	14,6%	14,9%	13,4%	14,6%
Sport	2,2%	1,1%	2,4%	4,1%	4,5%	6,1%	6,2%	6,5%	6,7%	5,8%	5,6%	6,2%	4,6%	4,8%	6,5%	6,9%	5,9%	6,0%	6,0%	4,5%	4,1%	4,1%	4,1%	3,7%	4,0%	3,8%	4,0%	4,2%	4,3%
Vasco	9,0%	2,2%	2,6%	2,7%	2,9%	2,4%	2,5%	2,4%	2,2%	2,0%	2,1%	2,7%	2,9%	3,0%	2,7%	2,4%	2,3%	2,2%	1,8%	2,2%	2,2%	2,4%	2,2%	2,2%	2,4%	2,3%	2,2%	2,5%	2,3%
Vitória	8,0%	3,3%	4,7%	10,2%	10,8%	10,1%	10,3%	9,2%	8,4%	8,3%	7,6%	5,7%	6,7%	5,5%	4,7%	5,4%	5,0%	4,9%	4,6%	4,4%	3,9%	3,7%	4,3%	4,0%	3,6%	3,5%	3,8%	3,8%	4,3%





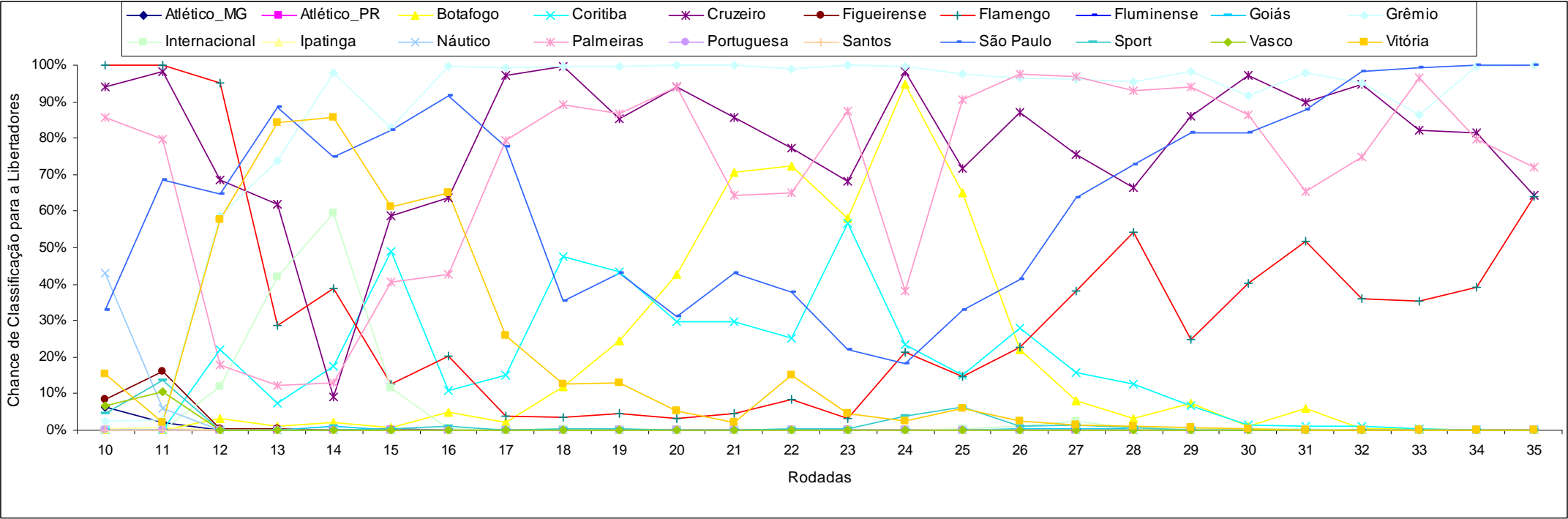
ANEXO H – Chances de ganhar uma vaga na Libertadores no Modelo 1



## ANEXO I – Chances de ganhar uma vaga na Libertadores no Modelo 2

Modelo 2	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
Atlético_MG	6,3%	2,2%	0,0%	0,1%	0,0%	0,2%	0,1%	0,0%	0,1%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Atlético_PR	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Botafogo	0,0%	0,0%	3,0%	1,0%	2,2%	0,8%	4,9%	2,0%	12,0%	24,5%	42,8%	70,5%	72,4%	58,2%	94,9%	65,1%	21,9%	8,0%	3,1%	7,3%	0,9%	5,9%	0,4%	0,3%	0,0%	0,0%
Coritiba	0,1%	0,0%	21,9%	7,4%	17,6%	49,1%	10,7%	14,9%	47,4%	43,3%	29,7%	29,7%	25,1%	56,5%	23,4%	15,1%	28,0%	15,8%	12,7%	6,6%	1,4%	1,2%	1,1%	0,2%	0,0%	0,0%
Cruzeiro	94,2%	98,2%	68,6%	61,9%	9,1%	58,8%	63,8%	97,1%	99,7%	85,2%	93,9%	85,8%	77,4%	68,1%	98,3%	71,7%	87,1%	75,5%	66,5%	85,9%	97,3%	89,7%	94,6%	82,3%	81,5%	64,2%
Figueirense	8,4%	16,0%	0,5%	0,2%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Flamengo	99,9%	100,0%	95,2%	28,5%	38,9%	12,5%	20,4%	3,8%	3,4%	4,5%	3,0%	4,7%	8,5%	3,1%	21,2%	14,7%	22,7%	38,2%	54,3%	25,0%	40,2%	51,9%	35,9%	35,2%	39,2%	63,9%
Fluminense	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Goiás	0,0%	0,0%	0,0%	0,0%	1,0%	0,0%	0,0%	0,0%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,3%	0,5%	0,2%	0,2%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Grêmio	2,6%	2,9%	58,7%	73,8%	97,8%	82,7%	99,7%	99,4%	99,5%	99,7%	100,0%	99,9%	98,8%	100,0%	99,8%	97,4%	96,5%	96,0%	95,3%	98,3%	91,7%	98,0%	95,0%	86,2%	99,5%	99,9%
Internacional	0,0%	0,0%	11,9%	42,1%	59,5%	11,7%	0,4%	0,0%	0,2%	0,0%	0,0%	0,1%	0,0%	0,0%	0,0%	0,4%	1,0%	2,4%	0,5%	0,4%	0,3%	0,3%	0,0%	0,0%	0,0%	0,0%
Ipatinga	0,2%	0,6%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Náutico	43,0%	5,9%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Palmeiras	85,6%	79,6%	17,9%	12,4%	13,1%	40,6%	42,6%	79,4%	89,2%	86,6%	94,1%	64,2%	64,9%	87,5%	38,0%	90,5%	97,6%	97,0%	92,9%	94,2%	86,4%	65,3%	74,9%	96,4%	79,8%	72,0%
Portuguesa	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Santos	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
São Paulo	33,0%	68,4%	64,6%	88,5%	75,0%	82,2%	91,5%	77,6%	35,3%	43,1%	31,1%	42,9%	37,7%	21,9%	18,3%	32,8%	41,3%	63,8%	72,7%	81,5%	81,6%	87,7%	98,1%	99,4%	100,0%	100,0%
Sport	4,7%	13,6%	0,1%	0,0%	0,0%	0,2%	1,0%	0,0%	0,5%	0,2%	0,0%	0,1%	0,3%	0,3%	3,8%	6,3%	1,1%	1,5%	0,8%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Vasco	6,7%	10,5%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Vitória	15,3%	2,1%	57,6%	84,1%	85,8%	61,2%	64,9%	25,8%	12,6%	12,8%	5,4%	2,1%	14,9%	4,4%	2,3%	6,0%	2,5%	1,3%	1,0%	0,6%	0,2%	0,0%	0,0%	0,0%	0,0%	0,0%

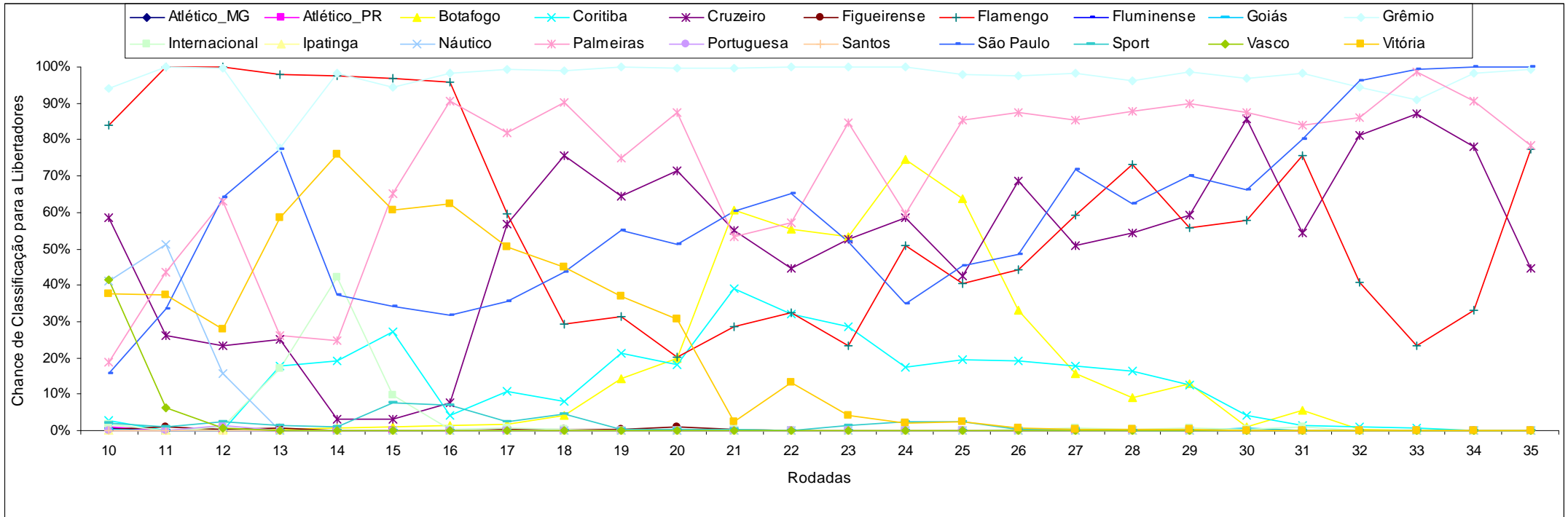
ANEXO J – Chances de ganhar uma vaga na Libertadores no Modelo 2





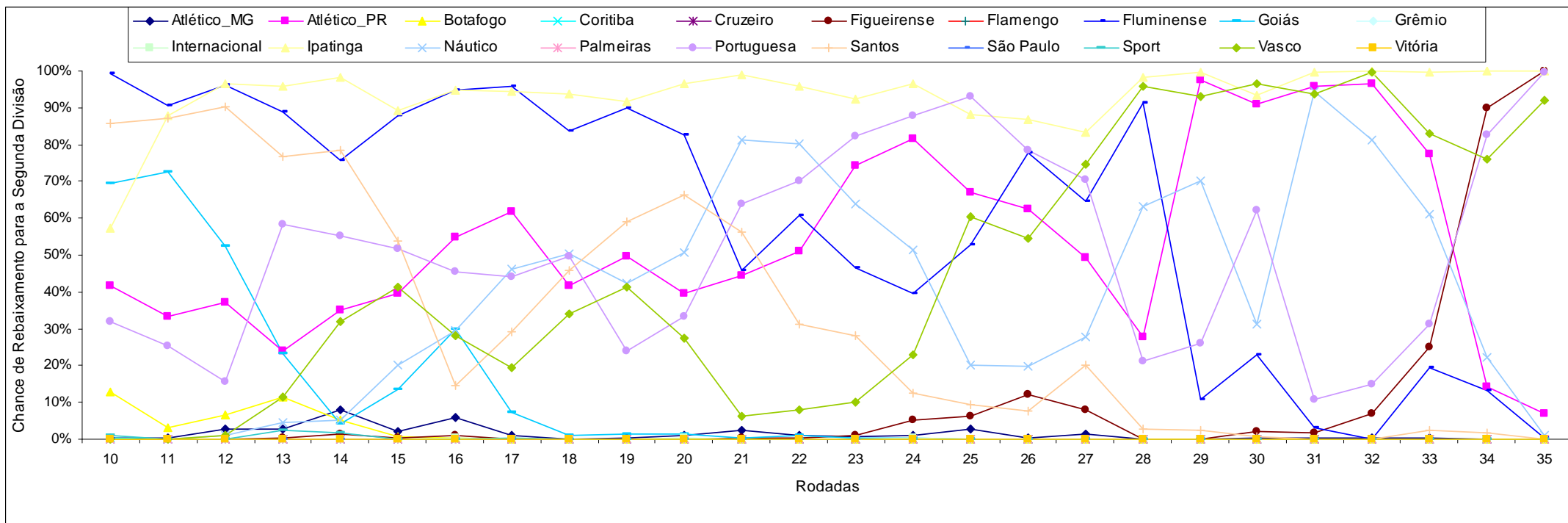


ANEXO L – Chances de ganhar uma vaga na Libertadores no Modelo 3



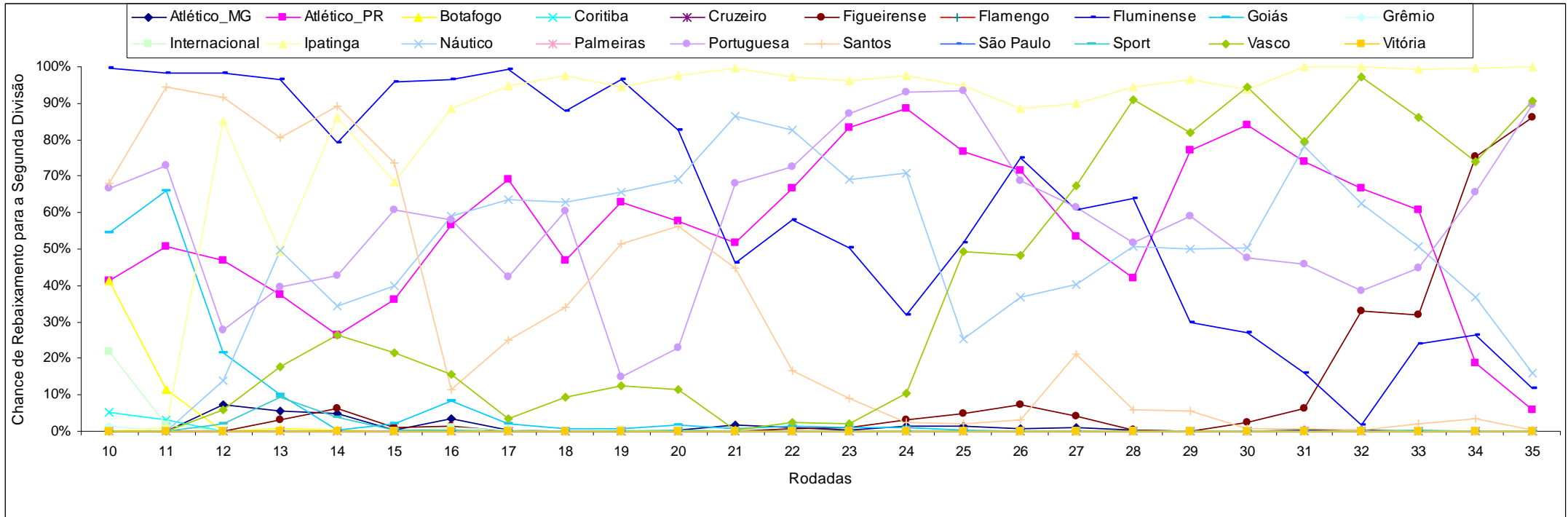


## ANEXO N – Chances de rebaixamento no Modelo 1



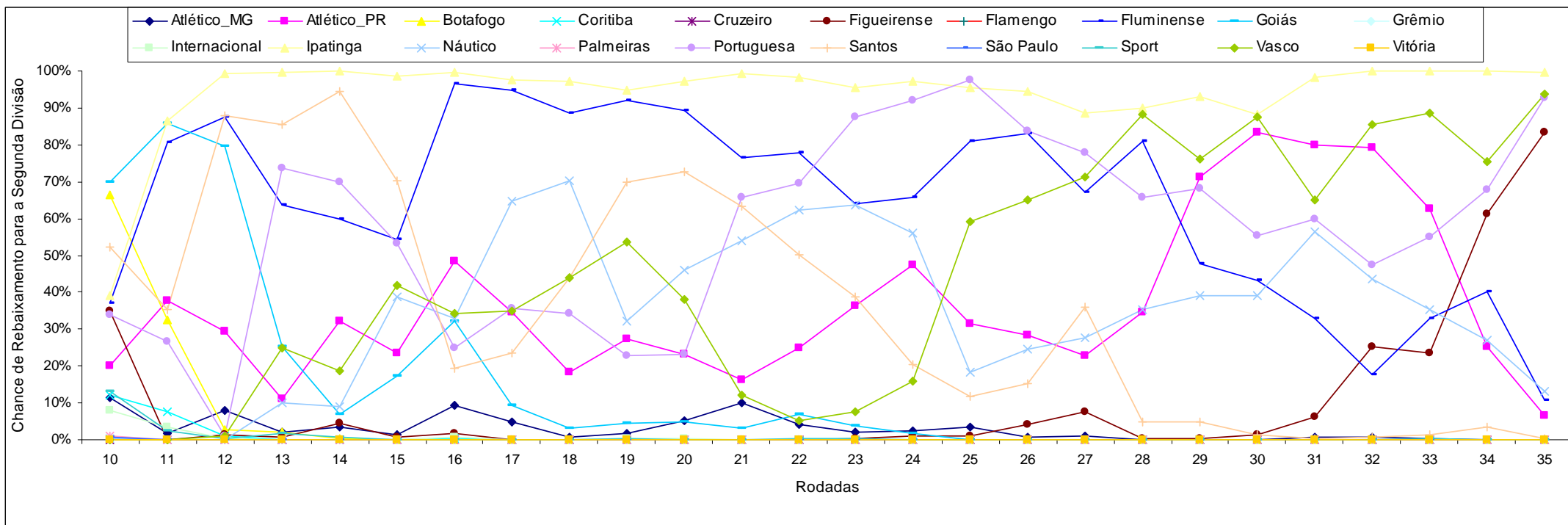


## ANEXO P – Chances de rebaixamento no Modelo 2





## ANEXO R – Chances de rebaixamento no Modelo 3





## ANEXO S – Coeficientes para o Modelo 1

Rodada	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
[Resultado = 1,00]	0.29	0.29	0.29	0.40	0.40	0.35	0.37	0.36	0.40	0.45	0.50	0.48	0.47	0.36	0.35	0.34	0.36	0.33	0.34	0.30	0.24	0.25	0.22	0.22	0.20	0.26	0.26	0.22	0.21
[Resultado = 2,00]	2.43	2.36	2.28	2.21	2.15	2.03	1.99	1.85	1.79	1.82	1.86	1.84	1.89	1.84	1.79	1.69	1.70	1.72	1.72	1.70	1.65	1.64	1.63	1.60	1.54	1.57	1.57	1.54	1.52
Atlético_MG	-0.02	-0.13	-0.20	-0.07	-0.20	-0.07	-0.23	-0.12	0.05	-0.07	-0.18	-0.25	-0.22	-0.18	-0.21	-0.27	-0.21	-0.25	-0.27	-0.16	-0.22	-0.25	-0.28	-0.20	-0.12	-0.09	-0.13	-0.16	-0.18
Atlético_PR	-0.91	-0.90	-0.78	-0.56	-0.63	-0.64	-0.71	-0.80	-0.71	-0.76	-0.73	-0.74	-0.80	-0.87	-0.91	-0.81	-0.75	-0.63	-0.68	-0.78	-0.79	-0.69	-0.65	-0.57	-0.46	-0.42	-0.36	-0.40	-0.33
Botafogo	-0.55	-0.38	-0.36	-0.36	-0.21	0.01	0.09	0.22	0.41	0.52	0.70	0.81	0.81	0.64	0.79	0.59	0.47	0.35	0.30	0.36	0.23	0.33	0.24	0.16	0.07	0.02	-0.04	-0.12	-0.03
Coritiba	0.04	0.07	0.41	0.28	0.30	0.48	0.26	0.44	0.58	0.64	0.59	0.63	0.54	0.59	0.45	0.37	0.48	0.37	0.44	0.42	0.36	0.31	0.35	0.28	0.20	0.17	0.21	0.12	0.07
Cruzeiro	0.88	0.96	1.02	0.83	0.46	0.75	0.79	1.02	1.08	0.88	0.95	0.84	0.71	0.60	0.71	0.53	0.65	0.56	0.62	0.65	0.73	0.62	0.68	0.60	0.61	0.53	0.59	0.52	0.54
Figueirense	0.22	0.46	0.49	0.18	-0.05	0.03	-0.01	0.14	-0.03	0.07	0.10	0.05	-0.10	-0.18	-0.30	-0.33	-0.42	-0.34	-0.23	-0.23	-0.30	-0.36	-0.46	-0.36	-0.45	-0.46	-0.41	-0.32	-0.27
Flamengo	2.22	2.26	1.76	1.35	1.24	0.94	0.80	0.56	0.42	0.47	0.48	0.59	0.65	0.50	0.63	0.53	0.55	0.61	0.71	0.57	0.65	0.70	0.71	0.60	0.67	0.73	0.67	0.60	0.53
Fluminense	-1.81	-1.37	-1.41	-1.05	-0.80	-0.97	-1.03	-1.17	-0.95	-1.06	-0.93	-0.73	-0.76	-0.60	-0.52	-0.58	-0.69	-0.61	-0.63	-0.51	-0.45	-0.35	-0.27	-0.37	-0.40	-0.36	-0.26	-0.18	-0.24
Goiás	-1.06	-1.12	-0.81	-0.51	-0.22	-0.46	-0.55	-0.43	-0.26	-0.30	-0.25	-0.16	-0.28	-0.19	-0.17	-0.01	0.05	0.11	0.13	0.09	0.15	0.03	0.00	0.07	0.01	0.06	0.08	0.13	0.06
Grêmio	0.76	0.79	0.82	0.92	1.17	0.95	1.20	1.23	1.21	1.36	1.42	1.30	1.23	1.24	1.19	1.00	0.94	0.86	0.90	0.92	0.81	0.86	0.80	0.69	0.79	0.83	0.75	0.80	0.82
Internacional	0.22	0.26	0.38	0.45	0.60	0.35	0.10	0.22	0.18	0.08	-0.03	0.08	0.03	-0.04	0.00	0.15	0.21	0.30	0.24	0.28	0.29	0.31	0.22	0.20	0.23	0.18	0.08	0.14	0.08
Ipatinga	-0.86	-1.23	-1.34	-1.13	-1.22	-0.97	-1.02	-1.12	-1.13	-1.04	-1.13	-1.20	-1.08	-0.89	-0.93	-0.78	-0.81	-0.74	-0.83	-0.85	-0.74	-0.83	-0.88	-0.77	-0.79	-0.73	-0.74	-0.80	-0.73
Náutico	0.57	0.14	-0.12	-0.23	-0.26	-0.51	-0.58	-0.75	-0.84	-0.75	-0.83	-0.99	-0.91	-0.72	-0.64	-0.48	-0.54	-0.52	-0.63	-0.64	-0.55	-0.61	-0.53	-0.46	-0.38	-0.31	-0.37	-0.33	-0.28
Palmeiras	0.92	0.71	0.79	0.48	0.49	0.66	0.77	0.89	0.94	0.84	0.92	0.78	0.80	0.88	0.69	0.81	0.86	0.83	0.85	0.85	0.77	0.68	0.71	0.80	0.70	0.66	0.67	0.67	0.57
Portuguesa	-0.77	-0.84	-0.58	-0.80	-0.73	-0.74	-0.65	-0.73	-0.80	-0.64	-0.71	-0.87	-0.90	-0.91	-0.95	-0.98	-0.86	-0.76	-0.79	-0.78	-0.67	-0.61	-0.56	-0.47	-0.52	-0.56	-0.58	-0.59	-0.61
Santos	-1.11	-1.15	-1.11	-0.80	-0.79	-0.64	-0.34	-0.52	-0.68	-0.75	-0.74	-0.67	-0.54	-0.42	-0.31	-0.27	-0.32	-0.40	-0.35	-0.37	-0.24	-0.18	-0.12	-0.20	-0.26	-0.20	-0.23	-0.20	-0.23
São Paulo	0.19	0.38	0.74	0.81	0.62	0.71	0.76	0.79	0.59	0.63	0.59	0.62	0.69	0.54	0.55	0.59	0.64	0.69	0.78	0.78	0.82	0.84	0.93	0.94	0.98	1.01	1.06	0.97	1.03
Sport	-0.18	0.22	0.08	-0.14	-0.04	0.16	0.24	0.12	0.16	0.10	-0.07	0.05	0.08	0.13	0.28	0.31	0.25	0.20	0.17	0.08	-0.01	-0.04	-0.09	-0.17	-0.10	-0.17	-0.12	-0.09	-0.04
Vasco	0.29	0.26	-0.07	-0.30	-0.48	-0.62	-0.47	-0.52	-0.66	-0.70	-0.59	-0.41	-0.41	-0.43	-0.52	-0.65	-0.68	-0.76	-0.89	-0.77	-0.85	-0.72	-0.77	-0.65	-0.58	-0.62	-0.67	-0.57	-0.64
Vitória	0.97	0.61	0.29	0.65	0.76	0.58	0.59	0.51	0.45	0.49	0.42	0.29	0.45	0.29	0.19	0.27	0.20	0.12	0.16	0.10	0.01	-0.03	-0.04	-0.12	-0.21	-0.27	-0.19	-0.19	-0.12

## ANEXO T – Coeficientes para o HOME ADVANTAGE do Modelo 2

Rodada	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
[Resultado = 1,00]	0.44	0.41	0.40	0.56	0.45	0.49	0.49	0.53	0.58	0.65	0.72	0.67	0.67	0.56	0.53	0.48	0.51	0.44	0.46	0.41	0.33	0.33	0.30	0.31	0.27	0.35	0.36	0.30	0.29
[Resultado = 2,00]	2.81	2.65	2.63	2.60	2.41	2.37	2.27	2.19	2.12	2.16	2.22	2.14	2.20	2.17	2.08	1.90	1.94	1.91	1.93	1.89	1.82	1.79	1.79	1.75	1.69	1.74	1.74	1.70	1.67
Atlético_MG	0.26	0.13	0.26	0.03	0.25	0.18	0.16	-0.12	0.30	-0.01	-0.18	-0.25	-0.26	-0.17	-0.13	-0.08	-0.02	0.02	0.09	0.17	0.04	-0.07	-0.06	0.03	0.27	0.32	0.31	0.32	0.32
Atlético_PR	0.21	-0.17	0.10	-0.19	0.01	-0.25	-0.32	-0.79	-0.79	-0.87	-0.98	-0.93	-1.04	-1.08	-1.05	-0.96	-0.86	-0.53	-0.47	-0.79	-0.73	-0.54	-0.39	-0.37	-0.03	-0.13	-0.01	-0.04	-0.02
Botafogo	-1.49	-0.67	-0.67	-0.65	-0.56	-0.38	-0.19	0.09	0.75	0.82	1.07	1.13	1.21	1.13	1.40	0.95	0.77	0.57	0.49	0.54	0.34	0.50	0.27	0.20	0.10	0.06	0.01	-0.10	0.00
Coritiba	1.21	1.03	1.02	1.18	0.64	1.39	0.83	1.11	1.38	1.44	1.47	1.42	1.38	1.43	0.99	0.81	0.93	0.78	0.76	0.77	0.62	0.56	0.58	0.50	0.43	0.42	0.41	0.21	0.21
Cruzeiro	1.71	1.51	1.46	1.81	0.93	1.51	1.34	2.01	2.13	1.77	1.73	1.70	1.67	1.50	1.75	1.05	1.25	1.08	1.19	1.20	1.35	1.17	1.08	0.93	0.89	0.94	0.99	0.96	0.93
Figueirense	-0.32	0.21	0.16	0.22	-0.15	0.14	0.01	0.35	-0.08	0.03	0.14	0.05	-0.13	-0.12	-0.29	-0.29	-0.38	-0.37	-0.15	-0.17	-0.28	-0.32	-0.53	-0.44	-0.59	-0.61	-0.56	-0.43	-0.40
Flamengo	2.25	2.15	1.96	1.19	1.34	1.05	1.13	0.69	0.74	0.68	0.58	0.75	0.82	0.69	0.97	0.79	0.81	0.84	1.02	0.91	0.97	1.02	0.97	0.87	0.94	1.02	0.99	0.94	0.92
Fluminense	-2.18	-1.87	-1.75	-1.98	-1.62	-2.01	-1.72	-2.30	-1.93	-2.03	-1.99	-1.35	-1.34	-1.18	-0.98	-0.88	-1.14	-0.85	-1.06	-0.77	-0.63	-0.54	-0.17	-0.48	-0.48	-0.47	-0.26	-0.18	-0.23
Goiás	-0.75	-0.90	-0.83	-1.02	-0.45	-1.03	-1.00	-1.09	-1.03	-1.06	-1.20	-1.10	-1.12	-1.10	-1.06	-0.73	-0.78	-0.79	-0.58	-0.59	-0.34	-0.62	-0.61	-0.48	-0.49	-0.43	-0.39	-0.33	-0.44
Grêmio	1.42	1.28	1.32	1.05	1.54	1.26	1.84	1.78	1.58	1.93	1.92	1.67	1.53	1.50	1.48	1.20	1.13	1.16	1.23	1.20	1.19	1.12	1.17	1.05	1.23	1.26	1.14	1.18	1.17
Internacional	-1.46	-0.76	-0.82	-0.23	-0.28	-0.47	-0.84	-0.41	-0.42	-0.60	-0.64	-0.35	-0.45	-0.39	-0.43	-0.05	-0.05	-0.07	-0.18	-0.11	-0.18	-0.02	-0.08	-0.03	-0.05	-0.03	-0.38	-0.35	-0.37
Ipatinga	-0.92	-1.53	-1.58	-1.50	-1.46	-1.42	-1.41	-1.58	-1.65	-1.64	-1.72	-1.80	-1.82	-1.68	-1.67	-1.52	-1.58	-1.57	-1.64	-1.63	-1.33	-1.48	-1.51	-1.36	-1.36	-1.39	-1.38	-1.41	-1.32
Náutico	1.37	0.80	0.76	0.30	0.46	-0.44	-0.52	-0.93	-0.93	-0.81	-0.82	-1.10	-0.98	-0.91	-0.89	-0.42	-0.52	-0.47	-0.67	-0.63	-0.51	-0.59	-0.59	-0.52	-0.45	-0.48	-0.57	-0.55	-0.51
Palmeiras	1.46	1.21	1.16	1.36	1.00	1.18	0.89	1.27	1.02	1.15	1.33	0.80	0.77	0.88	0.25	0.99	1.17	1.02	0.96	0.96	0.81	0.69	0.61	0.75	0.54	0.52	0.50	0.53	0.36
Portuguesa	-1.50	-1.32	-1.35	-1.20	-1.31	-1.03	-1.15	-0.94	-0.97	-0.58	-0.58	-0.87	-0.86	-0.95	-1.02	-0.83	-0.53	-0.58	-0.58	-0.52	-0.61	-0.48	-0.53	-0.35	-0.49	-0.54	-0.57	-0.61	-0.57
Santos	-0.72	-1.11	-1.02	-1.04	-0.78	-0.85	-0.35	-0.72	-0.89	-0.99	-0.98	-0.84	-0.81	-0.77	-0.63	-0.64	-0.69	-0.64	-0.64	-0.62	-0.36	-0.35	-0.22	-0.36	-0.40	-0.40	-0.39	-0.40	-0.44
São Paulo	1.07	0.77	1.22	1.08	1.03	1.06	1.08	1.06	0.77	0.78	0.89	0.82	0.78	0.70	0.65	0.60	0.68	0.82	1.00	0.92	1.01	0.91	1.27	1.22	1.29	1.31	1.43	1.31	1.38
Sport	-1.31	-0.67	-0.76	-0.86	-0.75	-0.29	-0.22	-0.29	-0.22	-0.30	-0.59	-0.48	-0.38	-0.37	0.03	0.00	-0.14	-0.25	-0.26	-0.32	-0.48	-0.36	-0.49	-0.55	-0.53	-0.53	-0.51	-0.44	-0.45
Vasco	-0.15	0.21	0.18	0.39	0.31	0.38	0.48	0.74	0.01	0.05	0.22	0.47	0.26	0.32	0.09	-0.45	-0.48	-0.54	-0.95	-0.81	-0.93	-0.68	-0.86	-0.60	-0.50	-0.57	-0.68	-0.52	-0.61
Vitória	-0.15	-0.29	-0.79	0.07	-0.15	0.03	-0.05	0.06	0.22	0.23	0.32	0.25	0.76	0.56	0.53	0.47	0.42	0.35	0.42	0.30	0.05	0.10	0.07	-0.02	-0.34	-0.26	-0.09	-0.10	0.06

## ANEXO U - Coeficientes para o Desempenho recente do Modelo 2

Rodada	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
Atlético_MG	2.77	1.25	-1.80	-0.65	-1.50	-0.43	-0.98	-0.10	-0.95	-0.21	-0.15	-0.19	-0.08	-0.28	-0.48	-0.60	-0.64	-0.80	-0.99	-0.89	-0.62	-0.51	-0.52	-0.60	-1.16	-1.26	-1.29	-1.36	-1.40
Atlético_PR	-6.88	-4.54	-2.49	-1.17	-1.63	-1.36	-1.25	-0.43	0.00	-0.10	0.49	0.31	0.42	0.37	0.25	0.42	0.27	-0.23	-0.42	0.12	-0.05	-0.27	-0.59	-0.41	-1.06	-0.61	-0.86	-0.89	-0.80
Botafogo	-0.24	-0.62	2.82	1.98	2.19	1.86	1.67	1.08	-0.99	-0.64	-1.35	-1.17	-1.33	-1.72	-2.12	-1.57	-1.39	-1.12	-1.10	-0.96	-0.66	-0.91	-0.32	-0.35	-0.03	-0.15	0.00	0.24	0.16
Coritiba	-5.84	-4.60	-0.90	-2.12	-0.07	-1.90	-1.09	-1.54	-2.60	-2.52	-2.84	-2.73	-2.65	-2.60	-1.62	-1.34	-1.53	-1.36	-1.17	-1.24	-0.99	-0.99	-0.94	-0.98	-0.99	-1.11	-0.85	-0.41	-0.56
Cruzeiro	1.58	0.72	-1.81	-2.74	-1.84	-2.39	-1.96	-3.30	-2.95	-2.69	-1.95	-2.13	-2.33	-2.26	-2.59	-1.38	-1.79	-1.58	-1.81	-1.60	-1.83	-1.68	-1.15	-1.03	-0.83	-1.14	-1.12	-1.20	-1.02
Figueirense	3.96	1.91	-0.73	-1.33	-1.02	-1.23	-0.89	-1.60	-0.27	-0.40	-0.27	-0.22	-0.07	-0.22	-0.07	-0.04	0.25	0.42	-0.01	-0.03	-0.04	-0.10	0.47	0.53	0.80	0.71	0.67	0.37	0.53
Flamengo	2.32	1.41	-2.37	-2.24	-2.34	-2.46	-2.62	-1.59	-1.62	-1.29	-0.98	-1.09	-1.03	-1.23	-1.81	-1.57	-1.29	-1.09	-1.51	-1.69	-1.84	-1.82	-1.59	-1.63	-1.91	-1.85	-1.93	-1.99	-2.17
Fluminense	-3.15	-0.65	1.21	2.38	2.45	3.01	2.08	3.34	3.19	2.96	3.23	1.95	1.73	1.76	1.43	0.98	1.52	0.96	1.44	0.88	0.68	0.67	-0.42	0.33	0.22	0.35	-0.12	-0.10	-0.13
Goiás	-4.09	-2.11	1.32	2.13	1.58	2.69	2.21	2.53	2.64	2.44	2.65	2.73	2.40	2.37	2.30	1.95	2.23	2.30	1.95	1.87	1.51	1.91	1.76	1.64	1.56	1.54	1.50	1.46	1.66
Grêmio	-6.68	-4.29	-1.51	-0.66	-0.94	-1.26	-2.28	-1.88	-1.25	-2.07	-1.69	-1.34	-1.16	-0.99	-0.96	-0.74	-0.73	-0.94	-1.03	-0.80	-1.06	-0.62	-1.02	-1.00	-1.16	-1.21	-1.07	-1.08	-1.01
Internacional	0.26	0.46	3.66	2.66	3.27	2.67	2.79	1.94	1.48	1.84	1.57	1.16	1.13	0.79	0.94	0.45	0.57	0.81	0.96	0.95	1.14	0.90	0.73	0.50	0.64	0.48	1.09	1.15	1.05
Ipatinga	4.79	5.06	2.10	2.91	1.80	2.14	1.75	1.82	1.58	1.95	1.80	1.89	2.27	2.31	2.22	2.16	2.12	2.16	2.22	2.10	1.66	1.89	1.75	1.59	1.44	1.69	1.63	1.65	1.57
Náutico	0.42	-0.84	-3.89	-3.37	-3.79	-1.39	-1.30	-0.23	-0.39	-0.58	-0.62	-0.14	-0.32	0.21	0.34	-0.68	-0.64	-0.73	-0.32	-0.43	-0.45	-0.33	0.09	0.14	0.15	0.46	0.59	0.61	0.65
Palmeiras	1.51	0.40	-1.73	-2.34	-1.56	-1.03	-0.49	-0.81	0.19	-0.34	-0.54	0.19	0.26	0.27	1.16	-0.25	-0.51	-0.28	-0.13	-0.13	0.01	0.10	0.38	0.26	0.59	0.46	0.53	0.47	0.72
Portuguesa	-1.09	-0.61	2.70	1.82	2.11	0.89	1.60	0.96	0.70	0.14	-0.04	0.17	0.07	0.33	0.37	-0.16	-0.67	-0.18	-0.26	-0.40	0.06	-0.13	0.03	-0.17	0.09	0.14	0.12	0.17	-0.03
Santos	-3.85	-2.18	-0.37	0.36	-0.88	-0.11	-0.12	0.34	0.94	1.03	1.12	0.82	1.20	1.55	1.54	1.65	1.52	1.02	1.14	1.00	0.65	0.75	0.50	0.64	0.69	0.86	0.66	0.71	0.70
São Paulo	2.00	2.15	-0.85	0.30	-0.25	0.05	0.08	-0.05	-0.12	0.16	-0.58	-0.23	-0.09	-0.23	-0.07	0.13	0.09	-0.13	-0.38	-0.18	-0.25	0.00	-0.74	-0.59	-0.70	-0.65	-0.85	-0.77	-0.90
Sport	5.22	3.79	1.25	1.13	0.94	0.23	0.68	0.45	0.75	0.80	1.41	1.51	1.52	1.59	0.99	1.10	1.20	1.33	1.33	1.24	1.41	0.95	1.13	1.03	1.09	0.96	1.00	0.89	1.01
Vasco	3.50	1.64	-1.46	-2.21	-2.46	-2.59	-2.43	-2.84	-1.22	-1.35	-1.42	-1.53	-1.23	-1.37	-1.12	-0.12	-0.09	-0.11	0.55	0.51	0.67	0.38	0.59	0.21	0.09	0.21	0.37	0.12	0.19
Vitória	3.50	1.64	4.85	3.15	3.97	2.63	2.56	1.94	0.88	0.87	0.14	0.03	-0.70	-0.67	-0.68	-0.42	-0.49	-0.44	-0.46	-0.32	0.01	-0.17	-0.11	-0.12	0.46	0.11	-0.07	-0.05	-0.23

## ANEXO V – Coeficientes para o HOME ADVANTAGE do Modelo 3

Rodada	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
[Resultado = 1,00]	0.26	0.30	0.23	0.23	0.24	0.15	0.22	0.26	0.33	0.37	0.44	0.38	0.32	0.27	0.26	0.27	0.27	0.27	0.28	0.25	0.16	0.17	0.18	0.20	0.17	0.24	0.25	0.20	0.19
[Resultado = 2,00]	2.84	2.89	2.67	2.34	2.26	2.13	2.06	1.95	1.94	1.95	2.02	1.94	1.93	1.95	1.91	1.81	1.80	1.85	1.80	1.76	1.68	1.68	1.70	1.70	1.62	1.66	1.65	1.60	1.59
Atlético_MG	-0.52	-0.65	-0.46	-0.34	-0.52	-0.51	-0.71	-0.68	-0.39	-0.48	-0.46	-0.63	-0.48	-0.44	-0.43	-0.49	-0.42	-0.33	-0.34	-0.14	-0.22	-0.29	-0.33	-0.23	-0.15	-0.13	-0.18	-0.24	-0.26
Atlético_PR	-0.78	-0.96	-0.55	-0.33	-0.44	-0.42	-0.53	-0.55	-0.47	-0.46	-0.42	-0.44	-0.56	-0.60	-0.68	-0.57	-0.51	-0.47	-0.62	-0.74	-0.74	-0.70	-0.67	-0.58	-0.46	-0.42	-0.35	-0.39	-0.32
Botafogo	-0.73	-1.05	-0.43	-0.33	-0.19	-0.03	-0.07	-0.02	0.03	0.19	0.24	0.51	0.54	0.55	0.70	0.58	0.53	0.35	0.25	0.28	0.18	0.28	0.18	0.08	0.02	-0.04	-0.06	-0.18	-0.02
Coritiba	-0.15	-0.34	-0.21	0.41	0.48	0.43	0.16	0.42	0.36	0.51	0.46	0.58	0.45	0.40	0.30	0.23	0.30	0.25	0.32	0.31	0.28	0.18	0.22	0.18	0.09	0.07	0.09	0.02	-0.02
Cruzeiro	1.29	1.09	1.16	1.24	0.63	0.71	0.71	0.89	1.03	1.00	1.08	0.89	0.85	0.84	0.97	0.78	0.91	0.73	0.75	0.83	0.91	0.86	0.92	0.89	0.90	0.86	0.91	0.85	0.87
Figueirense	-0.48	-0.47	-0.04	0.13	-0.07	0.13	0.13	0.07	0.05	0.26	0.36	0.27	0.03	-0.03	-0.14	-0.12	-0.23	-0.31	-0.20	-0.19	-0.27	-0.35	-0.42	-0.31	-0.40	-0.44	-0.39	-0.31	-0.24
Flamengo	1.91	3.81	2.72	1.33	1.38	1.14	0.96	0.78	0.63	0.58	0.53	0.62	0.68	0.55	0.62	0.54	0.61	0.63	0.61	0.41	0.41	0.49	0.46	0.33	0.40	0.49	0.42	0.37	0.29
Fluminense	-2.38	-2.29	-2.80	-2.47	-2.50	-2.41	-2.03	-2.32	-2.36	-2.52	-2.60	-2.35	-1.62	-1.42	-1.24	-1.32	-1.48	-1.16	-1.04	-0.89	-0.78	-0.78	-0.72	-0.79	-0.79	-0.70	-0.62	-0.41	-0.44
Goiás	-1.59	-1.46	-0.90	-0.66	-0.25	-0.48	-0.58	-0.41	-0.31	-0.34	-0.31	-0.26	-0.40	-0.33	-0.30	-0.12	-0.07	-0.06	0.22	0.21	0.30	0.25	0.20	0.24	0.15	0.19	0.20	0.19	0.12
Grêmio	2.07	2.64	1.92	0.99	1.55	1.39	1.46	1.65	1.64	1.75	1.80	1.75	1.70	1.67	1.57	1.33	1.28	1.25	1.20	1.30	1.21	1.19	0.99	0.86	0.95	0.97	0.89	0.95	0.94
Internacional	0.11	-0.02	0.34	0.58	0.88	0.57	0.26	0.37	0.29	0.17	0.11	0.22	0.17	0.07	0.11	0.20	0.30	0.32	0.30	0.30	0.28	0.37	0.25	0.19	0.24	0.18	0.09	0.13	0.09
Ipatinga	-1.09	-1.54	-1.92	-1.30	-1.35	-1.25	-1.29	-1.08	-1.11	-1.04	-1.15	-1.17	-1.15	-1.01	-1.04	-0.94	-0.92	-0.86	-0.89	-0.85	-0.73	-0.82	-0.92	-0.82	-0.85	-0.78	-0.78	-0.85	-0.81
Náutico	0.78	0.52	-0.06	-0.45	-0.45	-0.62	-0.59	-0.78	-0.88	-0.77	-0.83	-0.91	-0.85	-0.83	-0.76	-0.61	-0.62	-0.61	-0.79	-0.82	-0.71	-0.71	-0.57	-0.50	-0.42	-0.37	-0.40	-0.37	-0.29
Palmeiras	0.81	0.52	0.81	0.65	0.56	0.93	1.06	1.06	1.16	0.99	1.07	0.95	0.91	0.96	0.76	0.87	0.91	0.84	0.88	0.85	0.80	0.78	0.84	0.95	0.82	0.71	0.73	0.72	0.56
Portuguesa	-1.05	-1.29	-0.47	-0.85	-0.83	-0.74	-0.54	-0.78	-0.75	-0.70	-0.76	-0.96	-1.02	-0.99	-1.06	-1.09	-1.08	-0.89	-0.87	-0.89	-0.81	-0.76	-0.64	-0.51	-0.55	-0.62	-0.64	-0.64	-0.66
Santos	-1.19	-1.09	-1.55	-0.86	-1.15	-0.90	-0.48	-0.73	-0.86	-0.93	-0.92	-0.85	-0.80	-0.59	-0.52	-0.47	-0.55	-0.61	-0.51	-0.51	-0.39	-0.34	-0.34	-0.37	-0.43	-0.37	-0.42	-0.36	-0.42
São Paulo	0.85	1.24	1.55	1.23	0.94	0.84	0.86	0.89	0.85	0.91	0.92	0.94	0.88	0.79	0.74	0.81	0.89	0.94	0.95	0.92	0.94	0.99	1.13	1.17	1.22	1.25	1.27	1.15	1.23
Sport	-0.17	-0.17	0.05	0.19	0.30	0.55	0.55	0.63	0.67	0.53	0.51	0.59	0.25	0.26	0.54	0.57	0.42	0.40	0.40	0.11	0.00	-0.01	-0.01	-0.11	-0.05	-0.10	-0.04	-0.01	0.01
Vasco	1.22	0.54	0.13	-0.24	-0.14	-0.39	-0.36	-0.38	-0.46	-0.54	-0.45	-0.25	-0.20	-0.23	-0.34	-0.49	-0.51	-0.60	-0.78	-0.60	-0.63	-0.53	-0.64	-0.63	-0.55	-0.59	-0.63	-0.51	-0.60
Vitória	1.10	0.95	0.71	1.09	1.17	1.06	1.05	0.98	0.89	0.88	0.81	0.49	0.62	0.39	0.21	0.31	0.24	0.19	0.15	0.09	-0.05	-0.10	0.05	-0.03	-0.14	-0.18	-0.10	-0.10	0.00

## ANEXO Y – Coeficientes para o Desempenho Recente do Modelo 3

Rodada	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
Atlético_MG	-116.76	-14.62	-5.76	-4.61	-4.81	-6.34	-6.42	-5.13	-4.89	-3.79	-3.14	-2.92	-2.81	-3.25	-3.28	-3.43	-3.86	-3.15	-2.81	-1.74	-1.37	-0.96	-0.97	-0.81	-0.97	-1.05	-0.75	-1.03	-1.14
Atlético_PR	59.25	12.39	8.54	0.50	1.84	3.18	3.19	3.62	3.96	4.09	3.81	2.47	2.26	3.75	2.89	2.63	2.76	3.10	1.70	1.32	1.12	0.07	0.26	0.61	0.42	1.33	0.22	0.45	0.78
Botafogo	44.11	-8.65	-3.64	-5.32	-3.95	-4.77	-4.92	-4.70	-6.47	-5.98	-7.00	-6.42	-4.11	-4.12	-4.41	-4.64	-3.33	-3.84	-3.67	-3.51	-2.96	-2.57	-1.66	-1.60	-1.17	-1.23	-0.49	-0.74	0.06
Coritiba	50.39	1.61	-12.54	-1.58	-0.71	-3.60	-4.82	-2.19	-3.39	-2.92	-2.88	-1.75	-2.91	-3.10	-2.74	-2.71	-2.58	-2.77	-2.51	-2.03	-1.14	-1.65	-1.40	-0.99	-1.20	-1.28	-1.34	-0.90	-0.87
Cruzeiro	22.73	19.62	20.83	17.35	13.58	12.90	11.96	2.17	2.32	3.35	4.17	3.30	4.59	4.73	5.01	5.53	4.09	2.98	3.06	3.10	2.29	3.43	3.98	4.26	4.39	4.82	4.03	4.44	4.66
Figueirense	-104.89	-8.87	0.38	1.49	-0.32	0.03	1.62	-0.36	0.54	0.97	0.83	0.20	-0.62	-0.98	-0.36	-0.50	0.47	-0.32	0.09	0.44	0.43	0.32	0.59	1.00	0.90	0.47	0.19	-0.07	-0.42
Flamengo	-114.44	-4.56	1.55	-0.26	0.76	-0.26	-1.66	0.51	-1.01	-1.19	-1.00	-1.29	-1.41	-1.32	-1.80	-1.53	-1.50	-1.53	-1.79	-2.90	-2.90	-2.50	-1.77	-2.18	-2.28	-2.18	-1.92	-2.06	-2.20
Fluminense	42.15	-7.83	-13.67	-16.60	-16.03	-16.34	-9.22	-9.86	-10.50	-11.33	-13.07	-14.00	-9.90	-10.17	-9.73	-9.93	-10.85	-8.84	-4.66	-4.07	-4.06	-4.79	-5.20	-5.60	-5.61	-5.31	-5.47	-4.27	-4.54
Goiás	66.70	10.95	-0.51	-6.21	-4.72	-8.04	-6.34	-5.17	1.81	2.02	2.66	2.94	4.71	5.07	5.72	6.20	6.27	7.08	1.34	1.03	0.16	-1.36	-1.48	-1.82	-1.38	-1.51	-1.59	-1.11	-0.60
Grêmio	55.00	3.89	4.63	0.52	2.58	2.75	1.79	2.50	1.88	1.20	1.58	1.81	1.66	1.72	1.37	1.14	1.64	1.01	1.21	1.12	1.39	1.45	1.29	0.87	1.08	0.94	1.56	1.27	1.33
Internacional	61.10	11.20	8.97	1.20	3.75	6.29	4.58	5.58	4.86	3.98	4.70	3.11	2.55	2.45	2.19	2.55	3.11	3.39	3.48	3.72	3.84	3.13	2.14	1.55	1.63	1.69	2.03	1.72	1.80
Ipatinga	22.16	13.96	12.43	7.45	5.75	2.60	1.66	2.50	1.97	1.91	2.29	1.86	1.03	1.06	1.01	1.23	1.32	1.75	1.18	1.29	0.93	0.40	-0.32	-0.99	-1.10	-1.09	-0.97	-0.85	-0.66
Náutico	-122.74	-29.38	-17.65	-3.89	-3.79	0.43	-0.28	2.26	1.63	0.23	0.23	0.24	-0.31	-0.28	-0.45	-0.62	0.08	-0.47	-0.39	-0.59	-0.44	-0.23	0.60	0.45	0.59	0.45	0.98	0.51	0.63
Palmeiras	-115.79	-18.03	-8.02	-4.12	-4.79	-2.55	-1.00	-1.77	-0.51	0.13	-0.22	3.21	2.93	1.84	0.48	0.14	0.21	0.22	0.30	0.47	0.80	1.05	1.03	1.23	0.98	0.34	0.27	0.25	-1.04
Portuguesa	48.87	-1.81	-7.72	3.34	-0.19	-0.78	-2.62	-2.30	-3.05	-3.45	-3.29	-4.04	-4.08	-3.10	-3.42	-3.18	-5.11	-3.86	-3.19	-3.26	-3.54	-3.18	-2.22	-1.67	-1.33	-1.37	-1.22	-1.21	-1.27
Santos	50.65	-1.47	-1.71	-1.25	-2.13	-2.07	-1.01	-1.54	-1.97	-0.82	-0.56	-0.99	-1.65	-0.50	-0.87	-0.76	-0.74	-1.97	-0.83	-1.06	-1.57	-0.99	-0.44	-0.25	-0.18	-0.28	-0.10	-0.05	0.18
São Paulo	13.99	3.68	2.50	2.75	1.49	1.95	0.37	2.16	2.29	1.40	0.36	0.72	0.43	-1.08	-1.07	-1.65	-1.17	-0.83	-0.14	0.28	0.10	0.70	-0.75	-0.43	-0.85	-0.58	-0.72	-1.18	-1.67
Sport	-80.88	7.61	6.19	7.41	7.36	8.29	7.37	6.80	6.47	6.19	7.42	6.15	4.42	4.18	6.12	6.48	5.76	5.80	6.35	2.64	2.55	2.97	3.06	2.93	2.61	2.71	2.46	2.41	2.47
Vasco	59.20	5.16	8.35	5.70	6.09	4.37	3.21	2.57	2.90	2.35	1.61	1.71	-0.60	-1.07	-1.45	-1.54	-1.06	-1.50	-1.48	0.69	1.03	2.15	2.26	2.32	2.09	2.27	2.51	2.20	1.88
Vitória	59.20	5.16	-3.16	-3.86	-1.76	1.94	2.54	2.34	1.15	1.65	1.50	3.68	3.81	4.20	4.80	4.59	4.48	3.74	2.76	3.05	3.33	2.56	1.02	1.11	1.39	0.85	0.32	0.22	0.62

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)